Analysis of frozen ground in Finland:
affecting environmental factors, trends in northern Finland and applicability of satellite data

Aleksi Rimali

Master's thesis 791619S
Geography Research Unit
University of Oulu
March 2019

UNIVERSITY OF OULU

ABSTRACT

Faculty of Science

Appendix for M.Sc.-thesis
Maturity test for M.Sc.

Department:	Major subject:			
Geography Research Unit	**Geography**			
Author (Surname, forename): **Rimali Aleksi Vilhelm**	Student number:	**2263447**	Number of pages:	**106 + XI**

Title of the thesis:

Analysis of frozen ground: affecting environmental factors, trends in northern Finland and applicability of satellite data

Keywords: **Seasonal soil frost, Finland, 1981-2010, GAM, LME, Mann-Kendall, Sen, FT-ESDR, ERA Interim**

Abstract:

The warming climate will lead to major changes in cold regions in the future. These changes will be more rapid and more severe in high latitudes. This would eventually also affect the seasonal soil frost depth, which has a significant impact on water and energy cycle between atmosphere and soil surface in cold regions. The frozen ground is understood as the soil layer which freezes and thaws annually.

In this study, we investigate the seasonal soil frost depth in Finland 1981-2010 in open, forest and bog environments with three main aims: 1) To study which variables affect the overall thickness of soil frost layer and its changes most 2) To explore are there trends or major changes in the past 30 years on soil frost time series in northern Finland 3) To evaluate the applicability of satellite data against in-situ data in Finland. The main data in this study are frost tube in-situ measurements conducted by the Finnish Environmental Institute in 1981-2010. As a satellite data, we are using National Aeronautics and Space Administration's Earth System Data Record for Land Surface Freeze/Thaw State (FT-ESDR) and European Centre for Medium-Range Weather Forecast's (ECMWF) reanalysis model of soil temperature ERA-Interim datasets. For the first objective, we apply GAM (generalized additive model) and LME (linear mixed-effects model) statistical models in multivariate analysis. In the second objective, we are using the Mann-Kendall trend test and the Sen's slope estimate to conduct trend analysis. In the third objective, we evaluate the satellite-based measurement against in-situ observations with contingency tables.

Based on the multivariate analysis, the most statistically significant factors were air temperature, snow depth, precipitation and north coordinate. The interaction plots revealed that the effect of air temperature and snow depth to the maximum depth of soil frost is not a linear and varied in open, forest and bog environments. The yearly average of maximum depth of soil frost had decreased 2.12 cm/year on open, 2.75 cm/year on forest and 0.5 cm/year on bog sites in 1981-2010. The most distinct decreases were experienced in May in all three site types. The FT-ESDR and ERA Interim had the highest error rate percentages (avg. 68 % and 56 %) during shallow snow cover and soil frost depth. The accuracy increased steadily with the increasing soil frost and snow layer.

The study revealed that the seasonal soil frost depth has been decreasing between 1981 and 2010 in Finland. This study aimed to give more insight about the multidimensional process of frozen ground. Results can be applied in future research planning. The way to improve the current setting would require information about factors like soil moisture, groundwater, and extensive data from a longer period of time.

Further information:

Date: Oulu 6.3.2019

Contents

1. Introduction

In the next century, possible shifts in climate will be more rapid and more severe in high latitudes (IPCC 2014: 10). This so-called Arctic amplification is a result of solar radiation reflectioning from snow and ice. Changes in them would heat up the ground surface temporal and spatial range (Allison et al. 2001; Serreze et al. 2009; All About Frozen…2018). The environmental changes in the cold Polar Regions are in a key role in climate change effects globally. The Earth's average air temperature is rising, and this will eventually lead to changes especially in these regions which are more vulnerable to temperature fluctuations (French 2017: 37). Cold regions are also the main areas where cryosphere and all the phenomena that include it can be observed and studied. Cryosphere refers to the frozen environment where water is in the solid form as ice or snow. Its main phenomena are snow cover, sea ice, glaciers and frozen ground, which includes permafrost and seasonally frozen ground. On some of these phenomena, we can already observe predicted changes in the cryosphere in the past decades (Orradottir et al. 2008).

This study is focusing on seasonally frozen ground in Finland. Frozen ground has a significant impact on water and energy cycles between atmosphere and soil surface in cold regions. Changes in these areas have been seen as a strong indication of climate change (Allison et al. 2001; Sinhau & Cherkauer 2008). Today, the thickness of soil frost has decreased about 32 cm since 1930s (Frauenfeld & Zhang 2011). Recent studies have shown clearly that this trend is still continuing (Gregow et al. 2011a; IPCC 2014). The most distinct change in soil frost has been disruptions and shortening of the total length of frost season or even total disappearance from southernmost freezing areas. These changes are linked to rising air temperature, freezing index and snow depth (Frauenfeld & Zhang 2011: 19). According to World Meteorological Organization (WMO), the previous 2000-2009 decade was the warmest in history (Gregow et al. 2011b), and this has shortened annual freezing index, and on the other hand, increased thawing index. Both of these have been a commonly utilized factor in soil frost studies (Peng et al. 2013; Bilotta et al. 2015). In recent decades, permafrost studies have received clearly more attention than seasonally freezing ground, mainly because of the lack of reliable large-scale and long-term data, which have hindered soil frost studies (Zhang et al. 2005).

Major changes in permafrost and seasonally frozen ground could cause severe environmental, economic and social impacts. For example, forestry would suffer from shorter frost season, which would narrow the opportunities for logging during autumn. Unstable soils, like bogs, can support heavy logging machinery in a frozen state (Solantie 1998; Venäläinen et al. 2001a; Campbell et al. 2010). Frozen ground also anchors trees to the ground which reduces wintertime forest losses during heavy wind and heavy snow cover (Venäläinen et al. 2001b; Gregow 2013: 9). On the other hand, shorter frost season would also impair the effects of frost heaving and so could possibly reduce the annual maintaining costs of infrastructure like roads in the cold regions.

It is known that thawing ground releases greenhouse gases like carbon dioxide (CO_2), nitrous oxide (N_2O) and methane (CO_4) during thawing season, which is related to reduced plant uptake and microbial transformations (Öquist & Laudon 2008). It has been estimated that changes in annual snowpack and frozen ground may lead to increasing emissions from northern hardwood forests (Groffman et al. 2006). The wetlands of the boreal, subarctic and Arctic area have been estimated to contain about 35 % of the world terrestrial carbon pool (Gorham 1991). Shortening freezing season and longer thawing season during spring may lead to disturbances on the winter dynamics. Furthermore, these changes on soil-atmosphere interactions may also lead to transition on greenhouse fluxes in the cold regions (Oechel & Vourlitis 1994; Öquist & Laudon 2008).

The overall impact of the climate change to soil frost is not clear, and predictions of soil frost distribution should not be made based on climate nor factors as freezing index. If we want to understand climate change impacts to cold regions and frozen ground, studies considering seasonally frozen ground on a broader scale and long-term are needed. Recent studies have shown, that to understand the whole phenomena, we should take into account all geographical and geological aspects which could affect thermal and physical properties locally on soil frost depth (Yershov 2004: 352-355). Changes in soil frost extent, timing, duration and depth may have severe consequences, which can lead to a transition in whole climate and ecosystems (Zhang et al. 2004). In order to understand interdecadal variations in seasonally frozen ground depth and its distribution in regional and global scale, the more comprehensive long-term observations are needed (Luo 2017)

In Finland, soil frost is a common phenomenon and has been monitored by Finnish Environmental Institute (SYKE) since the beginning of the 20th century. As a Nordic country, the location of Finland on the high latitudes makes climate change and its effects on soil frost an essential topic (Venäläinen et al. 2001b). Soveri & Varjo (1977), Huttunen & Soveri (1993), Jylhä et al. (2009) and Gregow et al. (2011a, 2011b) have previously done research concerning soil frost in Finland. Jylhä et al. (2009) have concluded that annual mean temperature in Finland is rising approximately 4-7 °C and winter mean precipitation is projected to increase by 20-30 % (Jylhä et al. 2008). This is projected to lead to about 50 % smaller annual maximum to soil frost depth by 2100 (Venäläinen et al. 2001b: 63-72; Jylhä et al. 2008) and overall shorter frost season in Finland (Jylhä et al. 2008). To understand and make accurate projections about the future situation, we need to comprehend the factors, affecting soil frost depth. As Soveri and Huttunen (1993) and Venäläinen et al. (2001b) already pointed out in their study, there is a need and interest to continue soil frost research with more profound evaluation.

This study is aims to explore and compare the main factors which are affecting seasonally frozen ground in Finland. The three main questions are:

1) Which variables affect the overall thickness of soil frost layer and its changes most? We compare selected air temperature, snow depth, precipitation, solar radiation, freezing index, NAO-index and spatial parameter variables statistically. In addition, we compare the behaviour of these variables on open area, forest and bog. Based on previous studies, we hypothesize that air temperature, snow depth and freezing index are the most important factors.

2) Are there trends or major changes in the past 30 years on soil frost time series in northern Finland? We are using 12 frost tubes from northern Finland to search for differences in frost depths and trends in the area.
3) We evaluate the applicability of National Aeronautics and Space Administration's (NASA) Earth System Data Record for Land Surface Freeze/Thaw State (FT-ESDR) and European Centre for Medium-Range Weather Forecast's (ECMWF) reanalysis model of soil temperature ERA Interim datasets against the soil frost data in Finland.

2. Theory

2.1 Soil frost and freezing process

Seasonally frozen ground is a common cryosphere phenomenon in polar areas in both Southern and Northern Hemispheres. In the Northern Hemisphere, seasonal freezing occur about 60 % of the land area (All About Frozen…2018). The phenomenon can be understood simply as annual freezing of the ground, where a soil layer freezes and thaws seasonally (Williams & Smith 1989; Yershov 2004). The depth of this layer is dependent on local environmental and regional climatic factors (Williams & Smith 1989).

Freezing starts when the ground temperature falls below the freezing temperature of free liquid water. As a result, the water phase transition begins from liquid to solid. During the phase transition, free liquid water turns into solid ice, and the volume of freezing water increases by about 9 %. Usually, this volume increase does not cause any significant movement within the ground. The ground displacement due to soil freezing is called 'frost heaving'. Ground displacement is dependent on the groundwater level, soil type, and soil capillarity. The rate of freezing and total soil frost depth is dependent on the thermal-physical properties of the ground. Local factors, such as grain size, thermal conductivity, thermal gradient and amount of water in the ground as well as snow cover, freezing index and precipitation affect the process (Yershow 2004: 346-370). Ground freeze begins from the surface and advances to the deeper layers of the ground. The lower limit of the frozen soil layers is called 'frozen front'. Usually frozen front does not follow the 0 °C degree isotherm or the so-called 'cryofront'. Deeper parts of the soil cool slower mainly due to the thermal gradient of the ground and thermal conduction. Between the frozen front and the cryofront there is unfrozen or partly frozen ground which is called 'frozen fringe'. This transition zone of freezing soil contains water in liquid and frozen state. It is the warmest zone where water can be in solid form. Water is never completely pure in the ground, and impurities and dissolved salts lower the freezing point of water about 0.1 °C degrees. Liquid water contains so-called free energy which is consumed either to ground warming or freezing process during phase transition. Water releases latent heat to the ground at 0 °C degrees and heats up the whole soil. Latent heat slows ground cooling and causes so-called 'zero curtain effect' where grounds temperature stays constant until a sufficient amount of the latent heat is released to satisfy thermal gradient and thermal conductivity. During freezing the thaw water is being confined progressively in smaller space in cavities and free energy in the water decreases. This increases the pore water pressure in thaw cavities, and pores. Water starts to migrate to the freezing plane if free energy of water is higher on freezing plane than on unfrozen soil next to it. On freezing plane, thermal gradient causes now suction to freezing zone or so-called 'cryosuction', in which pore water migrates through frozen fringe from the unfrozen ground. This suction creates tension on the freezing plane and depending on its strength, the frozen front can either advance through pores creating pore ice on its way or stay stationary, and water migrates from underneath and leads to ice segregation. The amount of migrating water is also dependent on soil properties like grain size and soil capillarity. Fine-grained soil types have higher capillarity, so they can more easily maintain tension and cryosuction. This is more difficult, for example, on coarse-grained material like moraine matrix which contains larger clasts. Capillarity is understood as molecular forces between soil particle which, occur when the interface is confined. This force is stronger on smaller particles (Williams & Smith 1989: 5-10; Hentilä et al. 1994: 11-15; Thomas et al. 2009; French 2017: 66). Constant water supply to freezing plane leads to a creation of ice lens and frost heave (Thomas et al. 2009: 175). The freezing process and layers are presented with temperature in Figure 1.

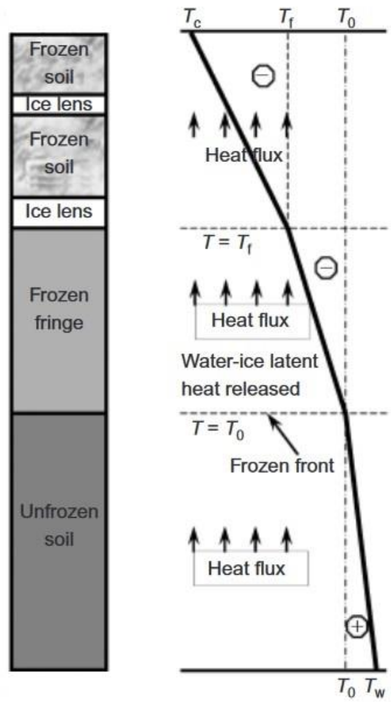


Figure 1. Schematic illustration of soil freezing process. Tc,Tw = Temperature at cold and warm ends. T_0 = Freezing point, T_f = Temperature at the base of ice lens (Thomas et al. 2009) Modified by author.

Freezing water creates different types of ground ice. This depends highly on soil properties, the amount of moisture and groundwater level. In the next section, we will briefly describe those ice types which are related to seasonally frozen ground. During these first cold weeks, 'needle ice' can occur on the top layers if ground contains a sufficient amount of moisture. These ice crystals develop during night frosts on the ground surface. Needle ice crystals may become as long as a couple of centimetres. They may push some minor particles and forest litter layer while growing and cause minor disruption on the surface (Williams & Smith 1989: 54; French 2017: 236). If the air temperature stays below freezing temperature, the heat release of the ground leads to deeper cooling and freezing in the soil. Deeper freezing creates pore ice, which forms when pore water freezes in ground pores, spaces, and cavities. This freezing substantially increases soil carrying capacity and term 'cement' ice is often used to describe this kind of ice. This is a common phenomenon on coarse-grained soil types which have low capillarity and absorption capacity, like gravel. The main ground ice type which creates frost heaving is ‘segregated ice’. ‘Pore ice’ forms in situ and does not require water migration (French 2017: 67). Segregated ice can be formed, if frozen front stays stationary and water migrates through frozen fringe to freezing plane. This creates an ice lens on the soil layer, and if the water supply is constant, the thickness of lenses can grow from a couple of millimetres to tens of meters. These layers or lenses are usually aligned horizontally. This type of ground ice is more common in fine-grained soil types, which have high capillarity and high absorption capacity (Williams & Smith 1989; French 2017). Soil freezing and frost heaving can cause annually major damage to infrastructure like roads and buildings (Venäläinen et al. 2001b). ‘Vein ice’ is formed when melt or rainwater intrudes in cracks and veins from ground surface and freezes on the ground. These cracks are the result of ground thermal-contraction. This process can create both horizontal and vertical ice lenses which separates it from segregated ice. ‘Intrusive ice’ can occur as a result of water intrusions and it usually requires water pressure. Good examples of this kind of ice are ‘pingos’, which have developed from pore water when water underneath has intruded under hydrostatic pressure toward the surface. Unlike segregated ice, intrusive ice requires high water pressure and this makes it possible to develop also in coarse-grained material (French 2017).

The climate of the area determines the occurrence of ground freezing. However the influence of local factors like microclimate of the area, vegetation, topography, soil properties, and snow cover have a tremendous effect on soil frost extent and its development (Williams & Smith 1989: 55). Even small obscure changes in local factors can lead to totally different conditions and soil frost depth in short distances. The multidimensionality is the reason why soil frost and its prediction based on climate has been seen as a very uncertain (French 2017: 15-55). The main factors which are affecting soil frost annual distribution and depth are presented in Figure 2.

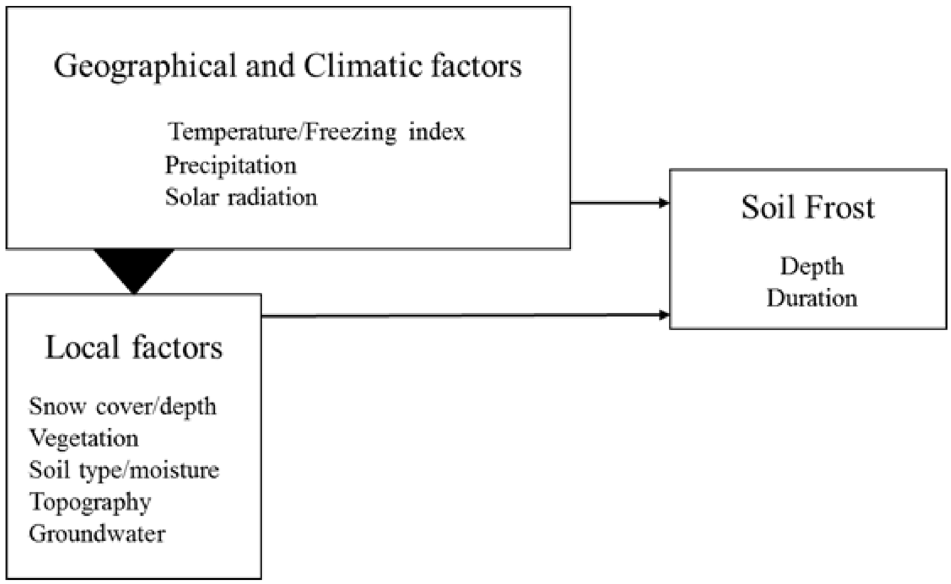


Figure 2. Environmental factors affecting soil frost (Mustonen 1986: 94-98; Williams & Smith 1989: 4-20; Seppälä 1999: 81-89; Venäläinen et al. 2001b; Salonen et al. 2002: 107; Yershov 2004: 346-364; Sutinen et al. 2008; Frauenfeld & Zhang 2011: 1-4; French 2017: 29-67).

2.2 Current extent, past changes and future predictions in the Northern Hemisphere

The seasonal soil frost distributional area consists of both Northern, Southern Hemisphere's polar areas and high altitude mountainous areas. In the Northern Hemisphere, this consist of about 55 million km^2, which is 60 % of the total land area. The southern boundary of seasonally frozen ground follows approximately 37 °N latitude in North America and 40 °N in Europe. In middle Asia, this boundary goes below 30 °N latitude duo to influence of high elevation areas and in East Asia, it follows 35 °N (Zhang et al. 2003).

It has been estimated that in Eurasia seasonally frozen ground depth has diminished 4.5 cm per decade and in total 31.9 cm between 1930 and 2000 in Eurasia. The decrease is linked to changes in winter's freezing index, mean air temperature and snow depth (Frauenfeld & Zhang 2011). The number of 'frost days' (Jylhä et al. 2008) and the duration of frost season have been decreasing (Henry 2008: 421-434; Bilotta et al. 2015). The number of frost days has been predicted to keep decreasing in the future, and, for example, frost season in Finland shortens (Venäläinen et al. 2001b: 63-72). In North-America, annual frequency of cycles of freeze- thaw are increasing in time wise (Henry 2008: 421-434).

In Finland, the annual maximum soil frost depth has been estimated to decrease on average by 50 % by 2100 (Venäläinen et al. 2001b). In southern Finland, the maximum soil frost depth in snow-free areas varies between 100 and 150 cm. In northern Finland, this maximum is about 100-300 cm (Venäläinen et al. 2001a). Frost season will shorten in southern Finland, but the average depth of soil frost might actually increase due to decreasing time of snow cover (Venäläinen et al. 2001b; Sutinen et al. 2008). Climate change is estimated to increase winter mean temperatures by 4-5 °C in Finland by the end of 21st century. Temperature and precipitation will increase in Scandinavia and this will lead to changes in vegetation and soil properties (Venäläinen et al. 2001b). In maritime climate zones, soil does not freeze deeply because of the high fluctuations on mean air temperature (Williams & Smith 1989: 15-18). Under continental climate, large variation between the temperature of different seasons leads to an increase in heat storage of the ground, but this requires that mean temperature is close to 0 °C (Williams & Smith 1989: 1-52).

2.3 Environmental factors

There are multiple factors in different scales which affect the occurrence of soil frost. Usually, some of these factors correlate with each other, which making the whole phenomenon difficult to analyse. Large-scale factors such as temperature, precipitation, and solar radiation are determined by geographical location and regional climate. On the other hand, local microclimate may differ significantly, affecting snow cover, soil moisture, and vegetation on the regional scale. Local factors like microclimate have a larger effect on local conditions and seasonal freezing occurrence than larger scale macroclimatic factors (Williams & Smith 1989: 59). Based on previous studies the main factors affecting the distribution, depth, and duration of soil freezing are air temperature, freezing index, snow cover, vegetation, soil type and soil moisture (Iijima et al. 2010). In the next section, we will describe these factors and their properties. Additionally, we present some other factors like precipitation, topography and solar radiation.

2.3.1 Local factors

2.3.1.1 Snow cover

Snow cover is as one of the main, if not the most important, factor determining soil freezing and thawing (Zhang 2005). It has a major influence on heat exchange between atmosphere and soil surface (Yershov 2004: 361). The thermal conductivity of snow is low, and it is varying for new dry snow between 0.12 and 0.46 W m^{-1} K^{-1}, which is five to ten times lower than mineral soil (Yershov 2004). For comparison, thermal conductivity of water and ice are 0.56 and 2.24 W m^{-1} K^{-1} respectively so the snow cover acts as an insulating layer between air and ground surface. Therefore, soil surface temperature can be much higher than air temperature. The phenomenon is called 'nival offset'. About 5-15 cm increase in snow depth raises the soil temperature approximately 1°C (Yershov 2004). With sufficient snow cover, it is possible that soil temperature stays above 0 °C even though the air temperature above is around −6 to −8 °C (Yershow 2004). The relationship between the depth of snow cover and its effect to soil temperature is not linear (Henry 2008: 430). Above 50-60 cm, increase in snow depth has a smaller effect on its thermal isolation (Williams & Smith 1989). The insulating effect of snow cover is at its best when the ground temperature is around 0 °C (Yershov 2004: 362). Snow has a major effect on surface energy exchange and this is determined by its timing, depth and duration (Zhang et al. 2005; Hirota et al. 2006; All About Frozen…2018). Snow thermal conductivity and insulating effect depend highly on its density.

During spring and summer, the nival offset effect turns around and snow cover keeps the ground temperature cooler than the air temperature. Increasing air temperature and solar radiation melt the snow cover and the insulating effect diminishes. Since wet snow has higher thermal conductivity than dry snow (Mustonen 1986: 94-98; Yershov 2004; Sutinen et al. 2008), melting can lead to a rapid temperature increase in the ground (Sutinen et al. 2008). Snow cover has a high albedo due which most of the solar radiation is reflected from the snow-covered area. High albedo and emissivity of snow layer surface keeps ground cooler (Williams & Smith 1989; Yershov 2004; Zhang et al. 2005). Albedo is determined by snow grain size and shape, surface roughness, and liquid water content and impurities, solar zenith angle and cloud conditions (Zhang et al. 2005: 1-3). However, longwave solar radiation can penetrate snow and heat up the ground. Approximately half of this radiation can penetrate 10 cm thick snow layer (Mustonen 1986: 94-98). Infiltration of snowmelt water through the soil surface in the ground can refreeze if the ground temperature is below 0 °C. This releases latent heat and increases snow and ground temperature (Zhang et al. 2005; Iwata et al. 2008).

Differences between seasonally frozen areas with and without snow cover can be substantial (Thomas et al. 2009: 173-184). In the end, the overall effect of snow cover on ground thermal regime depends on its timing, duration, accumulation and melting process of snow cover (Zhang et al. 2005). Properties of snow, like density and thickness, are affected by its interactions with vegetation, microclimate and geographical conditions of the area (Zhang et al. 2005). During the autumn, early thin snow cover accelerates ground cooling (Yershov 2004: 364-372) which typically lead to deeper soil frost compared to a situation where there is an early thick layer of snow present. On the other hand, early thick snow cover during autumn can prevent further freezing in the ground and insulates the soil (Iwata et al. 2008). Soil frost depth can influence spring flooding, because the frozen layer prevents soil infiltration (Bayard et al. 2005: 1). Snow cover and its influence on frozen soil play an important role when determining spring floods (Sutinen et al. 2008).

In Europe and North America, the southern line of seasonal snow cover follows approximately the seasonal soil frost area (Zhang et al. 2003). The extent of land area containing annual snow cover has decreased in Northern Hemisphere (Han et al. 2014) approximately by 10 % between 1972 and 2003 (Diaz et al. 2003). This decrease has been most severe during spring, and its rate increases with latitude due to larger albedo feedback

(Déry and Brown, 2007; Vaughan et al. 2013: 320). According to IPCC (2014) report, the mean annual extent of snow cover decreased approximately by 1.6 % in March and April since the mid-20th century. This decrease has been estimated in to continue in the future (Vaughan et al. 2013). However, there has been some uncertainty how the decreasing snow cover and its depth will affect soil frost. In some studies, it has been predicted that rising temperature and increasing precipitation would lead to decreasing snow cover and shorter frost season which would, eventually, reduce soil frost depth (Venäläinen et al. 2001b). On the other hand, some studies have concluded that increased precipitation would lead to thicker snow cover and shallow soil frost for example in northern Finland. Warmer winter temperatures are leading to shallower snow cover which on the other hand is predicted to cause deeper soil frost (Hardy et al. 2001; Campbell et al. 2010). In Finland, annual snow cover has been predicted to decrease due to rising temperature. This would probably lead to a shorter frost season in southern areas. In northern Finland increasing snow cover is estimated to reduce soil frost depth and all in all, conditions would be similar than previously in southern Finland (Venäläinen et al. 2001b).

2.3.1.2 Soil

Soil refers to ground layer and loose material on top of bedrock. This top layer consists of different soil types (Lehtinen et al. 1998; Kaivannaistietoa…2018). These properties have a direct effect on soil frost because they determine the rate and the extent of freezing and thawing in the ground. There are major differences between various soil type properties and other thermal factors which are affecting ground thermal regime and the formation of soil frost. The ground thermal regime is regulated by climate, surface and subsurface factors (Williams & Smith 1989: 19; Osterkamp 2007). In this section, we illustrate and emphasize these main soil properties which affect seasonally frozen ground.

Soils frost-susceptibility is determined by its thermal and hydrological properties, namely heat capacity, porosity, and thermal conductivity, and other soil conditions like moisture content (Mustonen 1986: 94-98). The first major thermal component is 'heat capacity', which is commonly evaluated either with the volumetric or mass-based method. Mass heat capacity is understood as the quantity of heat required to change the temperature of 1 kg of the substance by 1 K. This is expressed either on mass or volume basis. By multiplying mass heat capacity with substances density, we will get volumetric heat capacity C (J m^{-3} K^{-1}) (Williams & Smith 1989). Because soil commonly contains both mineral and organic material, the heat capacity is estimated with weighted average value using formula

$$C_s = X_m C_m + X_o C_o + X_w C_w \quad , \tag{1}$$

where X stands for the volume fraction of soil minerals (m), organic material (o) and water (w) (Williams & Smith 1989: 91).

The second essential thermal factor in ground freezing is 'thermal conductivity', which describes the rate of heat transfer by measuring the amount of heat flow through the unit area of the substance per unit time under a unit temperature gradient (W m^{-1} K^{-1}) (Williams & Smith 1989: 87; French 2017: 73). This varies between different grain size and soil materials like organic matter and minerals. Table 1 illustrates some of these materials and their thermal properties. The increasing water content in peat, sand and clay ground material increases the thermal conductivity. Increasing soil moisture and ice content raise its thermal conductivity which accelerates the freezing. However, liquid content in frozen soil makes evaluating thermal conductivity much more difficult, because it affects thermal-physical properties of ground material. This makes it especially hard to specify any clear temperature boundaries after ground temperature has passed 1 °C (Yershov 2004: 48-56). The greater amount of moisture in the ground means that phase change from liquid to ice will take a longer period of time, increases total amount of latent heat released to ground in this process (Yershow 2004: 358-359).

The so-called 'potential of soil material' is determined by its capillarity and absorption and these factors are also linked to its frost-susceptibility (Seppälä 1999). Soil grain size determines soils thermal conductivity, thermal gradient, hydraulic conductivity, and porosity. Porosity affects water filtration through the ground layer. Ground porosity and hydraulic conductivity can define ground ice type by regulating how much free water can filter into freezing zone. Soils which are considered frost-susceptible are usually coarse-grained. In general, coarse-grained and wet soil types freeze quicker and deeper compared to wet fine-grained soils (Soveri & Varjo 1977; Huttunen & Soveri 1993). On the other hand, in fine-grained soils, frost heaving is more common due to high capillarity, which is required for ice segregation and formation of ice lenses. In stable conditions, thermal gradient and thermal conductivity determines heat flow through the ground (Hinkel 1997: 15).

'The zero-annual amplitude' refers to the lower limit, where seasonal air temperature fluctuations can reach in the ground. This annual regime of climate is at around 10-20 m depth, and it is modified by snow cover and ground thermal properties (Williams & Smith 1989). Soil material works as a resistance which regulates heat flow. The rate of heat flow determines the amount of temperature change, and also the depth in which this thermal disturbance can reach. The Earth's interior acts as a heat source. This heat is flowing towards the surface, and therefore temperature is increasing with depth in the ground. This geothermal gradient refers to the distance in which temperature changes certain degrees with depth after moving one unit of length vertically. On the contrary, geothermic depth describes the distance at what ground temperature changes by 1 °C, which is typically 33 m. Both of these limit the soil frost depth. The inner heat flow from ground sets a lower limit for ground freezing depth.

Pore size and porosity of soil determines how much air and water the soil is capable to hold, and how much water is able to infiltrate to freezing zone. In frozen ground, thermal condition and changes in heat storage within a few degrees below the freezing point are controlled by latent heat effect, because of the temperature dependence of unfrozen water content (Williams & Smith 1989: 91).

Table 1. Thermal conductivity and mass heat capacity properties for different substances and soil material classifications (French 2017) Modified by author. Water masses for different soil types. Over 100 % = water mass greater than suspended matter (Salonen et al. 2002: 86) Modified by author.

Material	Thermal Conductivity (W m^{-1} K^{-1})	Mass Heat Capacity (kJ kg^{-1} K^{-1})
Air	0.02	1.00
Ice	2.23	2.09
Water	0.60	4.19
Organic matter	0.25	1.92
Clay minerals	2.92	0.90
Quartz	8.80	0.80
Peat		
Dry	0.05	
Saturated, unfrozen	0.50	
Saturated, frozen	2.00	
Snow		
Loose, new	0.086	
On ground	0.121	
Dense	0.340	
Rocks and materials		
Shale	1.5	
Granite	1.7-4.0	
Wood	0.12-0.16	

Material	Water mass (%)
Peat	500-1100
Clay	60-120
Silt	25-40
Sand	5-25
Gravel	5-15
Sandmoraine	10-15

Water bonds to soil particles with capillary, absorption and osmosis forces (Williams & Smith 1989). Of these, the first two are particularly essential to the freezing process. Soil frost is more common in slightly coarse materials like sand than completely fine-grained (Yershow 2004). Frost heaving is a result of ice segregation and is most common in fine-grained silt or silty moraine duo to their greater capillarity (Seppälä 1999).

The ground surface receives a different amount of solar radiation during the year which heats the ground. The amount of radiation varies between seasons, especially in cold regions. The ground heat flux is positive during the thawing season due to high solar radiation (Hinkel et al. 1997). This means that during the day ground absorbs more heat than it releases during the night. In winter this flux is negative (Hinkel et al. 1997: 1). The amount of stored heat can vary largely between different places even though the regional climate would be similar. In overall the ground is storing heat during summer months which is released on freezing season. This stored heat slows down the start of the freezing process and during winter all of the stored heat will be released (Williams & Smith 1989: 4-17). The rate of heat release and freezing depends on soil material and previously illustrated properties of soil and local climate factors (Williams & Smith 1989: 4-56; Hinkel 1997).

The freezing of moisture content reduces soil hydraulic conductivity, which might lead to an increased runoff due to decreased infiltration or even higher soil moisture content due to restricted drainage (Zhang et al. 2003; Zhang 2005: 19; Sinha & Cherkauer 2008). This frozen soil layer also prevents energy and water exchange between the atmosphere and the soil surface. Knowledge of the ground conditions during the freezing season would help in prediction surface of run-off and soil moisture (Zhang et al. 2003). The water content the soil is able to hold, varies between organic and mineral soils. Water content can be described with percentage mass of the water the soil contains. Table 1 shows water masses for different soil types. If the percentage is above 100 % water mass is higher than the mass of suspended material. In peat, soil frost penetrates faster than in mineral soil due to good thermal conductivity. Nevertheless, this is dependent on its moisture content. As shown in Table 1, the difference between thermal conductivity of dry and saturated peat is significant. Dry peat is a good insulator on the ground surface keeping the ground cool longer period time during the thawing season. On the other hand, during the autumn moist peat freezes and reaches greater depth faster than the mineral soils. The net effect is greater during thawing than in freezing season (Yershow 2004). In clay-rich soil types, water freezing temperature is usually clearly under 0 °C (between −2.5 °C and −3.5 °C).

The groundwater has a tremendous effect on its overhead ground thermal regime. On the other hand, frozen ground affects groundwater infiltration and the baseflow. The groundwater works as a heat source which releases heat energy towards ground surface. This means that groundwater level restricts soil frost lower limit on the ground. The principle is the same with rivers and lakes, which also heat up their surrounding areas. Water systems like rivers, lakes, and seas reduce temperature differences between different seasons (Mustonen 1986: 94-98).

2.3.1.3 Vegetation

The role of vegetation on soil freezing process is similar to snow cover. First, vegetation acts as an insulator like snow cover. This is called ‘vegetation offset’ because during the summer ground vegetation act as an insulating layer slowing ground warming. Its overall effect on ground temperature depends highly on the type, height, thickness, and closeness of vegetation layers (Mustonen 1986: 94-98; Williams & Smith 1989; Yershow 2004; Sutinen et al. 2008; Bilotta et al. 2014; French 2017). Compared to snow cover, vegetation cover effect to soil frost is much harder to evaluate, because it insulates ground both from cooling during winter and heating during summer (Yershow 2004: 363). This double insulating effect depends on summer and winter duration, continentality of climate, the depth of snow cover, the moisture content of the soil and several other factors (Yershow 2004: 364). Second, shrubs and small vegetation facilitate snow accumulation during early winter (French 2017: 44). Changes in these affect the annual amplitude of temperature fluctuations at the ground surface (Yershow 2004). Third, tree covers ability to intercept snow, leads to thinner snow cover in forest floor and deeper cooling of the ground (Hardy et al. 2001). However, the ground heat storage in forests is less than on open fields due to the canopy. The humus layer heat conductivity in dry state is weak and with canopy, this reduces heat storage in forests in thawing season. On the other hand, conductivity increases significantly on wet state during autumn compared to summer (Solantie 1998). During thawing season trees and smaller plants slows down soil frost thawing by reflecting and blocking solar radiation compared to bare ground (Yershow 2004: 364). This shading effect also slows down heat release and cooling of the ground in the beginning of the freezing season (Solantie 1998).

Vegetation influence to ground thermal regime is not straightforward and is determined by location and type. Grass level has an only a minor effect but for example, moss and lichen layer can have a major effect on ground temperature changes and freezing process. Particularly moss, lichen, and peat layers are highly effective insulator due to their low thermal conductivity. Already 2-3 cm moss-lichen layer reduces summer thermal sum by 66 % or even more but this is affected by current moisture content (Yershov 2004: 366). Dry moss-lichen thermal conductivity is about 0.1-0.7 W m^{-1} K^{-1}. This decreases by 66-50 in % on frozen state, making its insulating effect more significant during thawing than the freezing season. For example, about 15-20 cm layer of moss can reduce ground temperature around 5-6 °C. The thermal conductivity of peat in different states is shown in Table 1. Already 1 cm layer of peat can diminish ground average temperature by 0.5-1°C and just like moss, it is most effective in dry state during thawing season (Yershov 2004: 364-365). Even though soil surface temperature reaches over 20°C, the temperature may still remain close to freezing point around 20 cm below the surface due to the insulating effect of organic material (Hinkel 1997). Increasing solar radiation and temperature also increases, evaporation on the ground, which dry the surface moss and peat layer. The situation differs in autumn, when the moisture content increases due to lower evaporation and higher precipitation (Williams & Smith 1989). This leads to increased thermal conductivity, which accelerates ground cooling. In frozen state thermal conductivity increases even more and can lead to rapid freezing in the ground (Seppälä 1999). The overall insulating net-effect of peat layer is positive and thus it lowers mean annual ground temperature compared to areas which without peat cover (Williams & Smith 1989).

Even small changes in ground vegetation can correspond to 10-100 km shift on climate (Williams & Smith 1989), and removal of this ground level vegetation leads to greater average temperature fluctuations in the ground (Yershow 2004). Tree cover effect on ground thermal regime varies when moving from high to low latitudes in cold regions. Generally, tree cover and its overshadow reduces the amount and impact of solar radiation on the ground, and thus the heating effect of the sun, but this impact to ground thermal regime is determined by the total phytomass of surface (Yershow 2004: 364). Dense tree cover reduces snow cover thickness and therefore its insulating effect which leads to cooler ground. (Mustonen 1986: 94-98). However, the relationship with tree cover and ground temperature is not linear, and its effect becomes more unclear in the forest-tundra transition zone in cold areas. For example, Yershow (2004) stated that in measurements in the northern transition zone of forest-tundra in Russia, where vegetation consists of light forest and shrub, forest site mean annual ground temperature exceeded that of treeless site. This was caused by reduced turbulent heat exchange and thicker snow cover compared to treeless sites. However, in central and southern Russia where the forest is denser, these sites had lower temperature compared to treeless sites. This was a consequence of greatly diminished turbulent heat exchange, solar radiation and weaker winds, which decreased snow depth. The denser forest can hold more snow than sparse, but this differs between different tree types. For example, deciduous and pine forests' ability to hold snow is lower than spruce forests (Mustonen 1986: 94-98).

2.3.1.4 Topography

Changes in altitude lead to variation in of the several previously described factors. Site's topographical position defines local ground temperature regime and soil freezing depth (Yershow 2004; Shiklomanov 2012). Air temperature falls with altitude by about 0.4-0.6 °C per 100 causing a drop also in ground mean temperature (Yershow 2004: 359). Topographical factors like elevation, slopes, and aspect have a significant effect on areas air temperature, solar radiation, vegetation, soil composition and snow cover and in turn seasonal soil frost depth.

The orientation of slopes with respect to the cardinal point has a substantial effect for example, on the annual solar radiation balance of the ground. Southern and south-western slopes receive annually more solar energy than northern and north-eastern slopes, but this is significant only during summer. This means slower thawing in northern slopes compared to southern ones. Slopes and their steepness determine the incidence angle of solar rays which has a direct effect on the ground mean temperature and thawing depth. Slopes receive a different amount of solar radiation depending on their angle, and perpendicular slopes on 30 angle over sun are warming most during summer. However, overall low wintertime solar radiation in the cold regions is fairly equal in different slopes regardless of orientation. Soil properties and other local factors like grain size, humidity, snow cover and its thickness and vegetation, vary with altitude hindering the evaluation of soil frost depth. Even though factors like slope and aspect have an effect on ground thermal regime, this is only notable in thawing process. These are insignificant in freezing because cooling occurs almost uniformly (Mustonen 1986: 94-98, Yershow 2004: 360-362). Air temperature and ground temperature decrease northwards with latitudes because of increasing solar incidence angle and traveling distance of solar rays (Williams & Smith 1989: 59; National Snow…2010). Still, high elevation enables soil frost occurrence in areas where climate would otherwise be too warm, like the Himalayas and the Alps (Williams & Smith 1989).

2.3.2 Geographical and climatic factors

2.3.2.1 Air temperature

In large scale, the occurrence of seasonally frozen ground occurrence is determined by climate, but the local microclimate and other conditions can be seen as controlling factors. Climate defines mean air temperature of the region, and it is well known that air temperature affects to frost season length and soil frost depth (Bilotta et al. 2015; Luo et al. 2017). However, too often this seasonally frozen ground extent has been modelled only with the mean air temperature but this has usually led to errors, and is an inaccurate way to describe the whole phenomenon (Williams & Smith 1989). Based on the World Meteorological Organization's (WMO) guidance, climates should be statistically measured and present with 30 year periods. The word 'climate' refers to regional averaged weather in long period of time. Depending on seasonal variation, changes in air temperature affect ground temperature regime, fluctuations and in turn soil freezing process.

The freezing index is commonly used as one of the key explanatory factors in soil frost research (Williams & Smith 1989: 16; Gregow et al. 2011b: 11; Peng et al. 2013; Bilotta et al. 2015). The annual freezing index is calculated as a cumulative number of days when the mean of the daily air temperature is below 0 °C (All About Frozen…2018). On contrary, thawing index is the number of annual days when the mean air temperature is above 0 °C. However, there is variation about how this annual freezing index has been calculated. In some cases are used the sum of the daily mean air temperatures. For example, in Gregow et al. (2011b) study, this type of freezing index has been used in modelling of soil frost maximum depth on snowless areas with Stefan's equation

$$D = c\sqrt{F} \quad , \tag{2}$$

where *F* is the freezing index describing the sum of the mean air temperatures below 0 ºC with units: −ºC days. The coefficient *c* value vary depending on the soil type coefficient and for example, c = 5.49 cm (ºC days) $^{-0.5}$ value is used to describe sand soil thermal properties (Hentilä et al. 1994). However, this formula does not work on areas which are covered with snow most of the year due to the insulating effect of the snow cover (William & Smith 1989; Gregow et al. 2011b).

The global mean land and ocean surface temperature has risen on average by 0.7 °C 1850-2012, but overall this is estimated to be more severe and more rapid in the cold regions in the future (IPCC 2014). According to long-term measurements, last normal period 1981-2010 in Finland was about 0.4 °C warmer than the previous one 1971-2000, and almost 0.7 °C warmer than the period 1961-1990 (Heino & Hellsten 1983; Drebs 2002; Pirinen et al. 2012). Kellomäki et al. (2010) estimated that the annual mean air temperature would increase by 4.5 °C in Finland by the end of this century. The notable thing in these reports is that winter months were warming the most. Increasing winter air temperature could shorter snow cover time during spring, which in turn increases solar radiation emission. The albedo of bare ground is significantly lower compared to snow (Yershov 2004; Zhang et al. 2005). This increasing solar radiation emissions lead to warmer ground. The changes in these factors have a direct effect to freezing depth and freeze-thaw cycles on the ground (Frauenfeld et al. 2004; Henry 2007; Henry 2008).

2.3.2.2 Precipitation

Precipitation is the rainfall per surface area and time. The sun's radiation causes water to evaporate from lakes and seas and, depending on its movement in atmosphere, rains down as rain or snow returning to water cycle (Trenbert 2011; Strahler 2013: 94). Annual precipitation is controlled by regional climate conditions. This affect ground's moisture content and its thermal properties and in turn the freezing process. Rainfall, especially during autumn, plays a key role in the early freezing process of the ground. As was described in previous sections, higher moisture content improves ground thermal conductivity. In addition, wintertime mild periods, rainfall and snow melting during these times are affecting the development of frozen ground (Gregow et al. 2011b: 20).

IPCC (2014: 4) has estimated that in Northern Hemisphere the annual precipitation has increased over land areas in mid-latitudes during the 20th century. In Finland, annual precipitation is expected to increase 6-37 % by the end of this century (Jylhä et al. 2004, 2009; Ruosteenoja et al. 2005; Kellomäki et al. 2010). The winter precipitation is estimated to increase in Centre and Northern Europe (Jylhä et al. 2008: 452) which could mean overall thicker snow cover and in turn shallower soil frost depths in northern Finland.

2.3.2.3 Solar radiation

The Earth's 23.5 degrees tilted axel is creating seasonal changes in the planet. Rotation over this axel creates day and night changes. Duration of the day affects the amount of solar radiation and in turn extent of the frozen ground. The annual amount of solar radiation varies depending on the location and time of the year. Because the Earth is round, solar incidence angle changes and affects the amount of energy the surface is receiving. This flow rate of incoming solar radiation through atmosphere is called ‘insolation’, which is measured in units of watts per square meter (W/m^2). When this angle and distance to the surface increases, sun rays scatter more, diffuse and weaken when traveling through the atmosphere. Due to this, areas close to poles never receive as much radiation as the equator even though part of the year days are longer in these areas. However, the daily net radiation per m^2 during July in North Pole is higher than areas close to equator but the momentary effectiveness is smaller (Lean & Rind 1996; Solar radiation…2013, Strahler 2013; All About Frozen…2018). The solar radiation is also affected by local factors like topography, slope, aspect and shading tree cover (Williams & Smith 1989:66-69).

The solar radiation is affecting the ground thermal regime, especially during the thawing season, when net radiation is positive in the cold region. During winter, snow cover is reflecting most of the radiation even though long-wave radiation is penetrating snow layer and heating up the ground surface (Mustonen 1989: 94-98). This radiation from the sun is the primary factor controlling snow cover melting (French 2007: 239). The short wave radiation from the Sun is absorbed to the ground surface. This heated ground is in turn radiating long wave thermal radiation, which is blocked by greenhouse gases in the atmosphere and this leads to rising mean air temperature on Earth (Mitchell 1989; Strahler 2013; IPCC 2014).

2.3.2.4 North Atlantic Oscillation

The North Atlantic Oscillation (NAO) is the climate phenomenon which is associated with air temperature and precipitation fluctuations in Northern Hemisphere and is measured with NAO-index. This is represented either with positive or negative index value describing current atmospheric pressure differences between atmospheric low air pressure in Iceland and high in the Azores. Fluctuations between positive and negative values are occurring in the decadal scale. During the positive phase, pressure differences move cyclones from northeast to northwest into northern Europe, which leads to higher air temperature and increasing precipitation. In negative phase the situation is reversed, sub-polar low and subtropical high atmospheric pressures are lower than average, and this difference brings drier and colder air from continental Eurasia to northern Europe (Hurrell 2003; North Atlantic…2012).

This phenomenon is linked in some studies to the occurrence of seasonally frozen ground. For example, Frauenfeld and Zhang (2011) have investigated the correlation between soil freezing depth and NAO. In this comparison, they discovered moderate negative correlation (−0.58) between positive phase of NAO and freezing depth of the soil. There was also a strong negative (−0.89) correlation between negative phase and soil freezing depth. Correlation between the positive phase and soil freezing depth was measured in 70 year period from 1930 to 2000 and on the negative phase in 30 year period from 1970 to late 1990. Frauenfeld and Zhang (2011) found out that soil freezing depth especially during the 30 year period was likely driven by NAO-phase. Bojariu et al. (2008) studies showed that NAO-index influence in Eurasian cryosphere had increased recent decades significantly and linking autumns decreasing snow cover with following winter NAO value.

3. Research area: Finland

Finland's whole land area is used as a research area in this study, because of SYKE’s comprehensive frost tube network. Geographically Finland is located in Fennoscandia in northern Europe approximately between 60 and 70 N latitudes. Finland is part of southern and mid-boreal climate zone (Sutinen 2008) and northern parts are located at the zone of west winds at mid-latitudes. This is a transition zone of tropical and polar air masses, where weather types are changing rapidly especially during winter (Ilmasto-opas 2018).

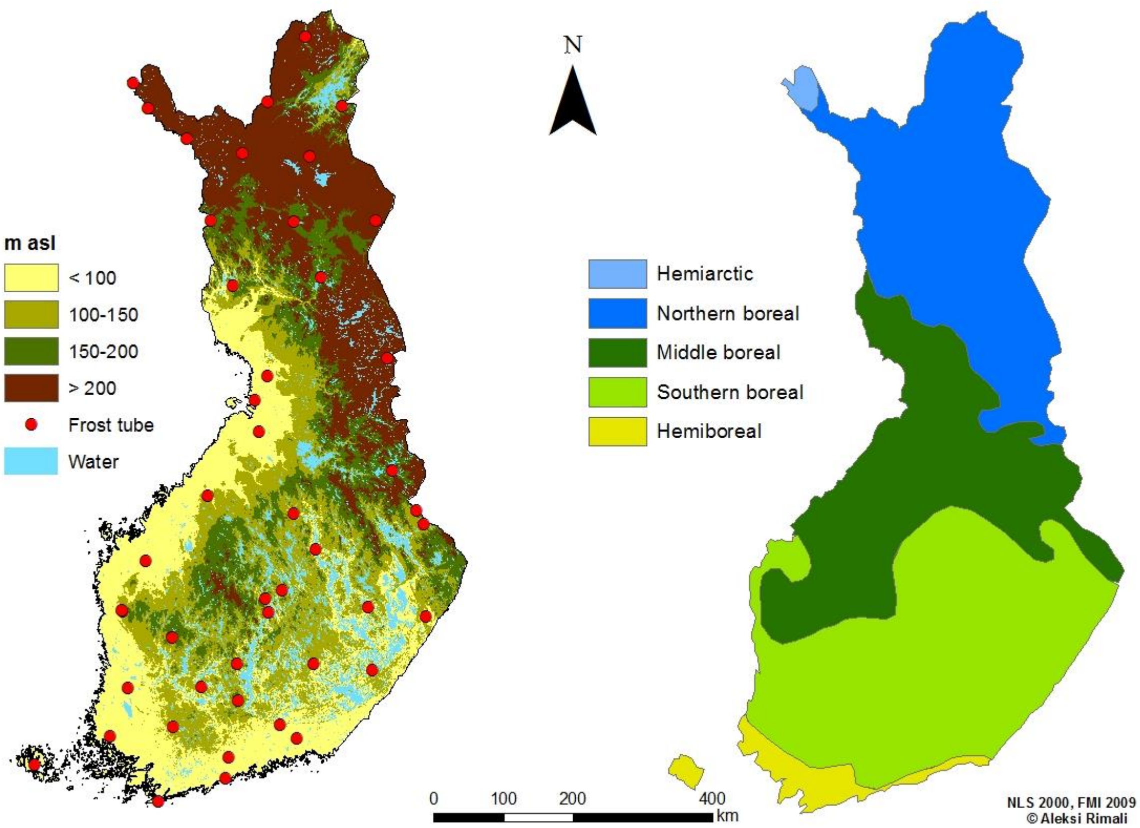


Figure 3. Frost tube locations and elevation map of Finland. The climate zone map of Finland (Kersalo & Pirinen 2009). Modified by author.

3.1 Geology

The bedrock in Finland is one of the oldest in the world and is part of the so-called Fennoscandian shield. Three major bedrock parts can be identified based on their origin. Eastern Finland and some parts of Lapland belong to Archean bedrock which has developed about 3100-2500 million years ago. Southern and middle parts of Finland formed about 1930-1800 million years ago and are so-called Neoproterozoic bedrock. There is also some minor parts of younger bedrock which have developed about 1650-1540 million years ago. The most significant part of these are Mesoproterozoic rapakivi granites in the southwest coast of Finland. After formation, there have not been any major changes in Finland's bedrock (Lehtinen et al. 1998).

The age of soils in Finland is relatively young, and most of the current soils have developed by erosion and accumulation of loose material during, and after the last glaciation. The most common soil type, moraine, was formed by ice sheets erosion forces. This have degraded bedrock and crushed material which in turn has accumulated ice melt water (Lehtinen et al. 1998). Almost 60 % of Finland's land area is known to be moraine (Ronkainen 2012: 8) which consist of fine-grained and coarse materials. This type of soil is classified as a diamicton, which means that its matrix includes smaller and bigger particles known as clasts (Salonen et al. 2002). Moraine type is typically either ablation till or ground moraine which cover almost half of the land are in Finland. The average moraine layer thickness is about four meters (Ronkainen 2012: 8). Gravel, sand formation, peat and clay deposits are also fairly common in Finland (Salonen et al. 2002).

There have been two major classification systems for soil types in Finland. In geotechnical soil classification, soil types are commonly classified by their geological origin, humus concentration, and grain composition. These soil types are divided into four main groups 1) Organic soil types 2) Fine-grained soil types 3) Coarse-grained soil types and 4) Moraine soil types. In Table 2 are all these soil types presented by their group and grain size.

Previously used engineering classification was in some parts different than geotechnical classification. There were some differences on grain size categorization and the sand class was divided into two separated classes, silt, and fine sand. These differences and soil type descriptions are illustrated in Appendix 1 & 2.

Table 2. Geotechnical soil type classification (Ronkainen 2012 modified by author).

Soil type class	Soil type	Soil type in Finnish	Shortening	Concentratio, weight-%			Grain size (mm)
				Clay	Fines	Gravel	
Organic soil types	Peat	Turve	Tv				
	Gyttja	Lieju	Lj				
Fine-grained soil types	Clay	Savi	Sa	>30			
	Silt	Siltti	Si	≤30	≥50	<5	≤0.06
Coarse grained soil-types	Sand	Hiekka	Hk		<50	≤50	>0.06…2
	Gravel	Sora	Sr		<5	>50	>2…60
Moraine soil types	Silt moraine	Silttimoreeni	SiMr		≥50	≥5	0.06
	Sand moraine	Hiekkamoreeni	HkMr		5…50	5…50	>0.06…2
	Gravel moraine	Soramoreeni	SrMr		≥5	≥5	>2

3.2 Climate and vegetation

Finland is part of so-called periglacial environments, which commonly refers to northern cold areas and occurrence of seasonally or perennially frozen ground. In Köppen-Geiger's climate classification Finland belongs to snow- and forest climate type with moist and cold winters. Typically in this zone, the warmest month's mean temperature is at least 10 °C and coldest −3 °C, and average annual precipitation is approximately 500-700 mm. The whole region is influenced by the presence of the Atlantic Ocean and Eurasian continental climate, which makes the regional climate consist of characteristics from both maritime and continental climates (Kersalo & Pirinen 2009: 7-10). Especially areas in west and south coast are affected greatly by the presence of the sea. In specific consideration, Finland can be divided into five climatic categories based on tree types starting from the south: 1) Hemiboreal 2) Southern boreal 3) Middle boreal 4) Northern boreal 5) Hemiarctic (Figure 3). The hemiboreal zone includes the southernmost coastline area and consist of oak and other hardwoods. The southern boreal zone is so-called maple and linden area and includes most of the areas on the south and middle Finland. The middle boreal zone is seen as a zone of mires where forest cover is more sporadic compared to South Finland. This area consist of parts from West Finland, northwest of Lapland and northernmost parts of Eastern Finland. Northern boreal is seen as an area which consists slow growing tree types, ‘aapa mires’ and includes most of northern Finland. The most northwest part of Lapland is part of Hemiarctic zone and here vegetation is mostly shrubs (Kersalo & Pirinen 2009: 7-10).

4. Data

4.1 Frost tubes and measurement sites

The main data for this study are frost tube in-situ measurements conducted by SYKE. The soil frost measurement network consists of 46 different sites covering the entire Finland. Frost tube measurement setup consists of an outer tube that is installed to the ground and an inner one filled with a mixture of distilled water and methylene. The outer tube is a cover and also in some sites as a scale for snow depth observations. The inner tube is used for soil frost observation. SYKE has conducted the frost tube observations with this method since 1970. The phase changes of inner tube substance – from liquid to solid and vice versa – follows the soil freezing and thawing changes around frost tube. Due to methylene, the substance inside the inner tube turns to blue when frozen. In liquid form the substance is colourless. The frost depth is read on the interface of blue solid and colourless liquid in the tube. The error margin is about ± 5 cm. (Orvomaa 2015; Orvomaa & Mäkinen 2015; Hydrologisen…2017).

Usually there are several frost tubes in a measurement site placed in three different locations in close vicinity. If possible one frost tube is installed in a location surrounded by a forest, one on open area, and one on bog. In every site, the local environment has been evaluated as well as the soil and forest types. Measurements are usually conducted on the 6^{th}, 16^{th} and 26^{th} days of the month in winter, and additional measurements on 1^{st}, 11^{th} and 21^{st} day of the month during thawing season if possible. There may be variations in measurement frequency in different years and sites. The measurements are continued until the soil is completely thawed (Orvomaa 2015; Orvomaa & Mäkinen 2015; Hydrologisen…2017).

4.2 Defining measurement period and sites

In this study, the data covers a time from 1981 to 2010, the 30-year period commonly used in climatic studies. In addition, the frost tube measurements prior to ’80s had more data gaps and were not as reliable and consistent as the selected period. In closer examination, several frost tube sites had 3-5 year long data gaps in 1981-2010. We excluded the widest gaps by setting 6 years as an upper limit. Every site surpassing this limit has excluded from our study. Based on this exclusion, 38 of all the 46 stations were selected in total. A stricter limit would have excluded too many stations out from dataset. It should be pointed out that the used variables and datasets differ in our three main parts of this study and these are described next in their own sections and also in Appendix 3.

4.2.1 Environmental variables affecting soil frost depth

As a response variable, we use the yearly average of soil frost monthly maximum depth. Here this variable is used to represent each winter in our data. In this study, the main explanatory variables are selected based on literature and the availability of data. Six main explanatory variables which are identified to have the most significant effect on the development and distribution of the seasonally frozen ground are air temperature, snow depth, precipitation, solar radiation, freezing index, and NAO index. For example DeGaetano et al. (1999) have used some of these variables in their prediction model to estimate the frozen ground extent in the US. For snow depth we used yearly average of monthly maximum, the same principle that was applied to frost depth. For snow cover we use monthly maximum values average for each winter season. Precipitation, air temperature, solar radiation, and NAO –index are yearly averages. We calculated freezing index for each winter from daily mean air temperatures records and its values are representing number of days with negative mean temperature.

For multivariate analysis, ten variables were included: 1) average of maximum snow depth, 2) mean air temperature, 3) mean solar radiation, 4) mean precipitation, 5) mean NAO index, 6) freezing index, 7) north and 8) east coordinates 9) year (season.id) and as a random factor 10) site ID variable.

Snow depth is measured from the frost tube locations at the same time with soil frost observations. The monthly NAO index value was provided by The United States Meteorological Institute on their website (North Atlantic…2012). This index represents atmospheric mass redistribution between the subtropical Atlantic and the Arctic (Hurrell et al. 2003: 1). It obtains values between 0 and 1 and one value summarizes our whole study area. The climate variables were extracted from Finnish Meteorological Institute's (FMI) climatological database which included monthly and daily measurements. The rasters have 10 km width cell resolution and each cell contains a monthly value. Based on frost tube locations we selected all those cells that contain one of our frost tube sites, and the selected raster cells were extracted with the values. As a result, monthly mean temperature, precipitation, and solar radiation are obtained for every soil frost observation site. For seasonal freezing index we used same scale raster but with daily mean temperature remote sensing dataset from FMI raster database. The seasonal freezing index was calculated for every winter from 1980 to 2011 by summing up the number of days in which mean air temperatures were lower than 0 °C (Hirota et al. 2006; National Snow…2012). The winter time period for the freezing index calculations was determined from September 1st to May 31st.

4.2.2 Soil frost annual maximum depth trend analysis in northern Finland 1981-2010

From northern Finland we selected 12 soil frost sites from the original 38 locations in Finland (Appendix 4). In the last SYKE's soil frost evaluation report, Finland was divided roughly into two main soil frost areas, Lapland and the rest of Finland (Orvomaa 2015). Due to this, sites close to the coastal zone and outside of northern Finland were not included in our study. By excluding coastal areas we try to minimize the effect of the sea to soil frost and keep study area uniform.

The maximum depth of soil frost was used in trend analysis instead of average depth. The soil frost monthly maximum depth should be a more stable and reliable statistic for describing soil frost in our long-term freezing season analysis. In our original data set, zero value represents no frost layer in the ground and NULL value that measurements haven't been conducted in that time. This could also lead to some errors when calculating monthly average values and therefore maximum values in this case should be more reliable.

4.2.3 Earth System Data Record for Land Surface Freeze/Thaw State and Interim Re-analysis of soil temperature

NASA's landscape freeze and thaw state evaluation product (Kim et al. 2011) uses passive microwave remote sensing observations. Its main purpose is to determine landscape's freeze/thaw status remotely. The algorithm uses daily radiometric brightness temperature measurements from Scanning Multichannel Microwave Radiometer (SMMR) and Special Sensor Microwave Imager (SSM/I) sensors at 37 GHz frequency. Both satellites are in Sun-synchronous and near-polar orbit passing the equator always at the same time. One full orbit provides one morning pass (AM) and one evening pass (PM). The number of orbits in one day is about 14 depending on the satellite. The FT-ESDR algorithm uses both AM and PM satellite overpass records and the final product represents daily global freeze/thaw information for the whole Northern Hemisphere. The record contains data for over 34 year's record period between 1979 and 2012. The FT-ESDR data is in categorized format consisting of four main classes for freeze/thaw state, presented in Table 3.

Table 3. FT-ESDR classification (Kim et al. 2011)

ID	Description
0	Frozen (AM/PM frozen)
1	Thawed (AM/PM thawed)
2	Transitional (AM frozen and PM thawed
3	Inverse Transitional (PM frozen and AM thawed)
251	No data
252	Non-cold constraint area, but unmasked
253	Masked
254	100 % open water
255	Fillvalue

The final product is provided in the Equal-Area Scalable Earth Grid (EASE-Grid) format with pixel size of 25 x 25 km. The latest version (3.0) of the data uses observations from SMMR and SSM/I satellites. The daily data are extracted from grid cells based on soil frost tube locations in Finland by interpolating the cells around each tube. Due to the complexity of setting limits for transitional class with records, we only use the first two classes from Table 3 to simplify the overall comparison. From all in-situ observations, we only use records from autumn and at the beginning of winter. The used annual period start on the 1st of September and ends on the 31st of December. Zhao et al. (2011) have pointed out that FT-ESDR estimates are inaccurate when the snow cover is melting. The increase in moisture content in the snow layer decreases the emission depth of microwave signal. As a result, the observation is not originating from the ground but only from snow layer.

In order to compare the measured frost depth data and categorized FT-ESDR data, which is a binary format, we have transformed the in-situ measurements to binary format as well. Because the most common measurement days during autumn were 6, 16 and 26 days of every month between September and December, we selected these dates from every year for analysis. These observations have been categorized into five classes based on frost and snow depth 1.) 0 cm, 2.) 1-5 cm, 3.) 6-10 cm, 4.) 11-20 cm and 5.) >20 cm. After this, all actual values are turned into a binary format where observed frost depths (>0 cm) represents FT-ESDR 0=frozen value and no-observations (0 cm) 1=thaw value. The snow depth was included as microwave remote sensing observations are less accurate with thick snow layers (Kim et al. 2011; Kim et al. 2017). Other values from Table 3 were excluded from the analysis.

The ERA-Interim is a global atmospheric reanalysis been produced by ECMWF. It contains information about several different variables and 4-dimensional variational analysis which is updated monthly. The forecast model is based on the 12-hour analysis window which uses ECMWF's Integrated Forecast System to produce reanalysis data. Data consist of 60 vertical levels from the surface up to 0.1 hPa. This product comprises analysis and forecast fields for several different atmospheric and surface parameters including soil temperature on four levels. Soil temperature parameter consists of forecast from four levels: 1) 0-0.07m, 2.) 0.07-0.28 m, 3.) 0.28-1.00 m and 4.) 1.00-2.89 meters depth in the ground. Spatial resolution for the final grid product is approximately 79 km and the coverage is global (Dee et al. 2011).

The level 1 (0.07 m) of ECMWF soil temperature data is used in this study from 1979 to 2010. These daily records are compared to SYKE's in-situ soil frost observations. In reality, the freezing point of ground may vary depending on the soil type but to simplify the analysis we assume that the soil is frozen when level 1 (0-0.07 m) soil temperature is estimated to be below 0 °C degrees. For comparison, we are again transforming temperature data and in-situ measurements with t same criteria's as with FT-ESDR data. The data are transformed into binary format by giving value 0 to temperatures below 0 °C and value 1 for temperatures above this. As we did with FT-ESDR, we are using in-situ snow depth measurements to classify in-situ soil frost measurements based on current snow thickness into five classes.

5. Methodology

5.1 Multivariate analysis

5.1.1 Basic principles of statistical modelling and study setting

In statistical modelling, the vital part is to identify what kind of data are used and how to fit the right type of model into it. When researching environmental phenomena and factors affecting it, the role of statistical modelling is to build a mathematical base for evaluation of relationships between response and explanatory variables. The aim is to determine whether the selected variables have sufficient explanation power over the response and is this power significant (Guisan 2002: 90-91).

In this study, we are aiming to investigate the overall magnitude of the effect of different explanatory variables to frozen ground maximum depth. As a response variable, we are using an annual average of monthly maximum thickness of soil frost layer (AVE.MAX.Frostlayer) consisting months from September to April. We are utilizing statistical tools which are based on linear regression model (LR) but are more flexible extensions of it and thereby suits better for our purpose. We decided to use the generalized additive model (GAM) in R statistical software, which allows possible nonlinearities in the analysis, to describe the relationship between explanatory variables and the response. Possible interactions between explanatory variables will be examined. The analysis will be implemented by first using GAM model to identify the direction of the effect of the explanatory variables and further investigate the overall magnitude by using linear mixed-effects model (LME) model after possibly transforming the explanatory variables as indicated by the GAM analysis. With GAM and LME, we will create a separate model for each environment type, open, forest and bog. In GAM models we are using 10 different variables in total (Table 4) and model evaluation is based on the estimated smooth functions. Air temperature and snow thickness are set under the same smooth term as well as north and east coordinates.

We include all the same variables in LME models as on GAM but in addition, we are also adding some quadratic variables on LME in order to model the nonlinear effects shown in the plots of the estimated smooth functions. Based on the literature, it is reasonable to expect that early thin snow layer has a positive effect on the formation of soil frost at the early stage. Hence we expect that the effect of air temperatures to soil frost is highly dependent on current snow depth as Venäläinen et al. (2001a) already have stated. Due to this, we are aiming to determine the major snow layer thresholds that modify the effect of air temperature on soil frost. Therefore mean air temperature and mean maximum of snow depth are included in our GAM models as an interaction term and identify possible thresholds from the interaction plots. Similar thresholds are examined from the plot of the effect of precipitation. Based on these thresholds, we advance to LME analyse with separate binominal variables from snow depth (SNOW.BIN.XX) and precipitation (Prec.BIN.XX), where 0 = below the limit and 1 = over the limit value XX. With these binominal variables, we can investigate more closely how they affect soil frost and how the effect changes after the observed limit. In addition to *gam()* function outputs, we calculate the statistics with *lme()* function to describe the effect of explanatory variables on soil frost layer thickness. We calculate season’s average for every variable based on monthly data from each year

In Table 4 are presented further description about used variables in GAM and LME models, including interaction terms for air temperature, snow cover, and precipitation. The variables are evaluated based on their magnitude of the regression coefficient estimate and p-value. The variables which did not show sufficient evidence of statistical significance (p-value >0.05) were removed. The snow and precipitation terms (Prec.BIN.XX and SNOW.BIN.XX) were kept in the model if the interaction terms were statistically significant.

Table 4. Used variables in GAM and LME models. The XX is representing the chosen snow depth and precipitation value.

Model	GAM	LME	Description
Response	AVE.MAX.Frostlayer	AVE.MAX.Frostlayer	Maximum thickness of soil frost layer
Explanatory variable	AVE.MAX.Snow	AVE.MAX.Snow	Average of maximum snow layer thickness
	AVE.NAO	AVE.NAO	Average NAO-index
	AVE.Prec	AVE.Prec	Average precipitation
	AVE.Rad	AVE.Rad	Average solar radiation
	AVE.Temp	AVE.Temp	Average air temperature
	Freezing.Index	Freezing.Index	Cumulative summation for days with average air temperature below 0 °C
	season.id	season.id	Year ID variable
	coord.east	coord.east	East coordinate
	coord.north	coord.north	North coordinate
		SNOW.BIN.XX	Binomial snow depth variable
		Snow.BIN.XX;AVE.Temp	Air temperature interaction
		Snow.BIN.XX;AVE.MAX.Snow	Snow depth interaction term
		Prec.BIN.XX	Binomial precipitation variable
		Prec.BIN.XX;AVE.Prec	Precipitation interaction term
Random factor	site	site	Frost tube site ID

We are including a categorized variable *site* as a random factor to both models which separates different frost tube based on the location. This helps us to take into account differences between frost tube sites all over of Finland. As a time variable, we added *season.id* variable which separates different freezing seasons 1981-2010. Also spatial variables coordinate north and east are under the same smooth term. These east and north coordinates are KKJ Finland uniform coordinate system also known as YKJ which unit is in meters (m) (National Land…2010).

In multivariate analysis, a common problem is that some variables might use the same type of information to explain the distribution of the response variable. This phenomenon is known as multicollinearity (Vatcheva et al. 2016) and this was tested between all variables with Spearman's correlation test (Elliot & Woodward 2007: 2-33) separately for open areas (O), forests (F) and bogs (B) in freezing season analysis. The Spearman's correlation coefficient matrices are presented in Appendix 5. Correlation coefficient over 0.7 is usually indicating too high collinearity between explanatory variables. The most common cut-offs for multicollinearity in research, are values over 0.5 or 0.8 (Vatcheva et al. 2016). We decided to use the value of 0.7 as an upper limit for identifying multicollinearity between variables in this study. Spearman correlation coefficient values are interpreted same way as Pearson's where values are between 0 and 1 where 0 means zero and 1 perfect correlation between variables (Elliot & Woodward 2007: 2-33). In general, absolute values between 0.1-0.3 are indicating no correlation, 0.3-0.5 weak correlation, 0.5-0.7 moderate, 0.7-0.9 strong correlation (Hinkle et al. 2003). These cut-offs might vary depending on discipline.

Before implementing GAM model in three environment types, we are evaluating correlations of selected explanatory variables with Spearman's correlation test as described previously. From here the variables are dropped out based on possible multicollinearity problem where >0.7 correlations are considered as a limit.

5.1.2 Linear Mixed-Effects Model

Linear mixed-effects model is a popular statistical analysis method which can explain the relationship between response and covariates in case of repeated measurements. The major advantage of the model is its flexibility to handle within-subject correlations. The basic linear mixed-effect model can be written as

$$y_i = X_i\beta + Z_ib_i + \varepsilon_i \qquad (3)$$

where y_i is observed responses with a n_i -dimensional vector. X_i and Z_i are (n_i x p) and (n_i x q) matrices. These terms include β fixed effects p-dimensional vector and the random term with b_i random effects q-dimensional vector. Both of these terms can contain explanatory variables and ε_i is within-subject errors. The b_i are assumed to be independent with $N(0, \Psi)$ and ε_i is assumed to independent with $N(0, \Delta_i)$ distribution (Pinheiro et al. 2001; Zuur et al. 2009: 101-139; Bondell et al. 2010).

5.1.3 Generalized Additive Model

Generalized additive model is a more flexible semi-parametric extension of linear model. The term semi-parametric comes the fact that even though GAM can model predictors non-parametrically, distribution of response variable specification is still needed (Guisan 2002). Its strength is that it does not choose beforehand any parametric form for the response or explanatory variables and so prejudge response curves shape. GAM uses the 'smooth' function for explanatory variables (Luoto & Hjort 2005). The fitness of the model is controlled by changing degrees of freedom and so changing level of smoothness in the functions. GAM's basic structure is same as generalized linear models but with added smooth functions 'f_i':

$$g(\mu_i) = a + f_1(\beta_1 * x_1) + f_2(\beta_2 * x_2) + \ldots + f_k(\beta_k * x_k) \qquad (5)$$

In our analysis, we are using Gaussian distribution as default family for all the models and as link function we use identity function i.e $g(\mu_i) = \mu_i$ (Hastie & Tibshirani 1999).

The GAM model is using by default Generalized Cross Validation (GCV) for smoothing parameter estimation for model terms. However, in our model construction, the degrees of freedom are adjusted based on model estimated smooth function plots. If the estimates of the smooth functions are indicating that degrees of freedom need to be tuned, we are setting the number by manually. As a site is a random factor, the degrees of freedom are not constrained.

5.2 Analysis of northern Finland soil frost and its possible changes in 1981-2010.

Our main interest in these northern tubes concerns the yearly maximum soil frost and snow depths, and possible changes in these over three decade time period. In addition, we are interested in the starting day of soil freezing and thawing and the total length of frost season in different locations and environment types. Because we are interested in phenomena which are not normally distributed, we decide not to use a simple linear model for analysis. Possible trends in time series will be analysed first with Mann-Kendall trend test and with Sen's slope estimate to determine the possible monotonic trend and its total magnitude. The analysis is carried out with the Excel template (MAKESENS) created by FMI (Salmi et al. 2002). We also take a closer look in the discussion section about different soil types and how these have possibly affected soil frost depths.

The starting date for the frost season during autumn and formation of ice in the soil is determined by finding the first day when soil frost depth was over 0. For the thawing season, we used top layer depth values of frost and searched the first day of the spring when frost layer have been started to melt from the top.

5.2.1 Mann-Kendall's trend test

The commonly used method to detect and analyse possible changes in environmental and climatic time series is trend analysis. Mann-Kendall trend test is a widely used statistical tool to research trend and its direction from time series (Déry & Brown 2007; Sinhau & Cherkauer 2008; Longobardi & Villani 2010). This test has been utilized for example, on snow cover extent (Déry & Brown 2007) and rainfall time series analyses (Longobardi & Villani 2010). This test is non-parametric and it does not assume normally distributed data. The test contains two parts depending on the number of observations, but overall test follows model:

$$x_i = f(t_i) + \varepsilon_i \tag{6}$$

where f(t) is continuously decreasing or increasing function of time and the ε_i residuals with zero mean are assumed to be from the same distribution. We are testing null hypothesis H_0 = no trend i.e. our observations x_i are randomly ordered against the alternative hypothesis H_1 = increasing/decreasing monotonic trend. In case where n < 10 the absolute value of *S* is compared directly to the theoretical distribution of S derived by Mann-Kendall. Here two-tailed test is used in four different significance level α 0.001, 0.05 and 0.1. The H_0 is rejected in favour of H_1 at certain probability level if the absolute value of S is equal or exceeds a specified value $S_{\alpha/2}$, where $S_{\alpha/2}$ is the smallest S which has probability less than α/2 to appear in case of no trend. A positive (negative) value of S is indicating increasing (decreasing) trend. (Gilbert 1987). The *S* test statistic is calculated with equation (7):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_i)$$

and (7)

$$sgn(x_j - x_i) = \begin{cases} 1 & if x_j - x_i > 0 \\ 0 & if x_j - x_i = 0 \\ -1 & if x_j - x_i < 0 \end{cases}.$$

The sign of the test statistic S indicates the possible direction of the trend in the time series. When *n* is 10 or higher, we can use a normal approximation as follows. First, we can compute the variance of S with an equation (8).

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{q} t_p(t-1)(2t_p+5)\right] \quad (8)$$

Where q is the number of tied groups and t_p is the number of data values in the p^{th} group.

The variance VAR(S) and test S values are used to compute test statistic Z as

$$Z = \begin{cases} \frac{S-1}{\sqrt{VAR(S)}} \; if \; S > 0 \\ 0 \; if \; S = 0 \\ \frac{S+1}{\sqrt{VAR(S)}} \; if \; S < 0 \end{cases} \qquad (9)$$

As in test S, the value from test Z is used to identify the presence and the direction of the statistically significant trend in time series. The positive (negative) Z value indicates increasing (decreasing) trend. The tests validity is reduced if there are several equal values and n is close to 10. The monotonic trend is tested (a two-tailed test) at α level of significance where the H_0 is rejected if the absolute value of Z is greater than $Z_{1-2/\alpha}$. Here $Z_{1-2/\alpha}$ obtained from standard normal cumulative distribution tables. The p-values are tested in three significance levels α 0.001, 0.05 and 0.1 (Gilbert 1987; Salmi et al. 2002; Longobardi & Villani 2010).

5.2.2 Sen's method

Sen's nonparametric test estimates the true slope of the existing trend as change per unit time. The test assumes that the trend is linear and follows the next equation (10):

$$f(t) = Qt + B \qquad (10)$$

Here *Q* is the slope and *B* constant. First, we compute slopes for all data value pairs to get the slope estimate *Q* in the previous equation.

$$Q_i = \frac{x_j - x_k}{j-k} \qquad (11)$$

where j > k.

When there are *n* values x_j in the time series, and we can get as many as N = n(n−1)/2 slope estimates *Qi* and Sen's estimator of slope is the median of them. The *N* values are from ranked smallest to largest and the Sen's slope estimator is computed either with

$$Q = Q_{[(N+1)/2)]}, if \; N \; is \; odd \qquad (12)$$

or

$$Q = \frac{1}{2}\left(Q_{[N/2]} + Q_{[(N+2)/2]}\right), if\ N\ is\ even \tag{13}$$

Slope estimate's 100(1−α) % two-sided confidence interval is determined by the nonparametric technique which based on normal distribution (Gilbert 1987; Salmi et al. 2002).

5.3 Earth System Data Record for Land Surface Freeze/Thaw State and ERA-Interim

Due to the high resolution of FT-ESDR and ERA-Interim raster datasets, the daily cell values are interpolated around each frost tube site in Finland. We conduct analysis with simple cross tabulation method for the comparison of SYKE in-situ measurements with NASA's and ECMWF's data. Here dichotomic in-situ observations counts and frequencies are compared with two datasets separately with different snow depths. FT-ESDR and ERA-Interim daily values are sorted by dates when in-situ observations are measured. From FT-ESDR data we just use 0=frozen and 1=thaw values and ERA-Interim are sorted based on these dates. These observations are finally divided based on in-situ observations into a 5x5 table. Snow depth measurements are divided into four class (0-5, 6-10, 11-20, >20 cm) and soil frost depth into five class (0, 1-5, 6-10, 11-20, >20 cm).

5.3.1 Cross tabulation

Contingency tables, also known as cross tabulation, is a popular statistical method for qualitative research for example in biological and social science. In the case where we have categorical variables or binomial data, cross-classification of these tables makes it easy to represent these counts and to compare them. The table can be formed with two or more variables where each of them are representing one dimension. In these tables, several variables are illustrated in counts and frequencies of observations (Fienberg 1994).

In this study, we are using two-way cross tabulation method to simply illustrate the validity of NASA's FT-ESDR landscape freeze/thaw state and ECMWF's ERA-Interim soil temperature reanalysis with SYKE's in-situ observations during a different depth of soil frost and snow cover. The satellite observations will be divided into five classes based on in-situ snow and soil frost measurements: 1) 0 cm, 2) 1-5 cm, 3) 6-10 cm, 4) 11-20 cm and 5) >20 cm. With this 5x5 contingency table we are able to determine how satellite observations accuracy is varying with snow and soil frost depth conditions and how much differences there are between these two remote sensing models. The results are presented as error rate percentages of total observations for making the comparison of the accuracies in different snow and soil frost classes easier.

6. Results

6.1 Environmental and climatic variables affecting soil frost depth in Finland

6.1.1 Descriptive statistics and correlations

Table 5 presents basic summary statistics of the variables, which were used in freezing season's analysis. It should be pointed out here that all frost tubes locations did not include all three types of environmental frost tube sites. The Enontekiö Kilpisjärvi's measurement site does not include tube on open and neither Kaisaniemi in the forest. Bog tube was missing in six locations, Anjala, Utsjoki, Pälkäne, Mietoinen, Jokioinen, and Kaisaniemi. In all these cases, sites were included in data but values marked as missing value (NULL). We also included east and north coordinates as one variable under same smooth term in the model to get a comprehensive picture of the spatial distribution of soil frost maximum depth in Finland.

Table 5. Response variable (orange) and six explanatory variables used in freezing season GAM analysis.

OPEN AREAS	Unit	Data	Min	Max	Mean	Median	SD	N	NA
AVE.MAX.Frostlayer	cm	in-situ	0	149.9	28.5	19.2	27.9	565	537
AVE.MAX.Snow	cm	in-situ	0	75.5	29.7	28.2	15.9	977	126
AVE.Prec	mm	grid	15.7	79.2	43.7	44	10.7	1102	0
AVE.Rad	W/m^2	grid	2484	7420	4074	4034	664.3	1102	0
AVE.NAO	index value		−0.3	0.5	0	0	0.18	1102	0
AVE.Temp	°C	grid	−9.9	4.0	−2.7	−2.5	3.1	1102	0
Freezing.Index	cumulative sum	grid	88.3	2343.8	1154	1150.7	403.4	1102	0
FORESTS	**Unit**	**Data**	**Min**	**Max**	**Mean**	**Median**	**SD**	**N**	**NA**
AVE.MAX.Frostlayer	cm	in-situ	0	134.6	28.4	17.0	28.1	972	130
AVE.MAX.Snow	cm	in-situ	0	116	35	32.5	18.3	889	213
AVE.Prec	mm	grid	15.7	79.2	43.7	44	10.7	1102	0
AVE.Rad	W/m^2	grid	2484	7420	4074	4034	664.3	1102	0
AVE.NAO	index value		−0.3	0.5	0	0	0.2	1102	0
AVE.Temp	°C	grid	−9.9	4.0	−2.7	−2.5	3.1	1102	0
Freezing.Index	cumulative sum	grid	88.3	2348.8	1154.6	1150.7	403.4	1102	0
BOGS	**Unit**	**Data**	**Min**	**Max**	**Mean**	**Median**	**SD**	**N**	**NA**
AVE.MAX.Frostlayer	cm	in-situ	0	67.5	13	11	10.9	714	388
AVE.MAX.Snow	cm	in-situ	0	106	42.9	44.5	23.7	714	388
AVE.Prec	mm	grid	15.7	79.2	43.7	44	10.7	1102	0
AVE.Rad	W/m^2	grid	2484	7420	4074	4034	664.2	1102	0
AVE.NAO	index value		−0.3	0.5	0	0	0.2	1102	0
AVE.Temp	°C	grid	−9.9	4.0	−2.7	−2.5	3.1	1102	0
Freezing.Index	cumulative sum	grid	88	2343	1156	1151	403.5	1102	0

From the first part of Appendix 5, we can observe that air temperature, solar radiation and precipitation are affecting negatively (>0.4) to frost layer thickness in open areas and forests. Precipitation (AVE_Prec: O: −0.575, p <0.01; F: −0.584, p <0.01) air temperature (AVE_Temp: O: −0.639, p <0.01; F: −0.761, p <0.01) and solar radiation (AVE.Rad: O: −0.448, p <0.01; F: −0.508, p <0.01) are also affecting negatively to soil frost depth. Snow layers thickness correlation was moderate with solar radiation (O: −0.639, p <0.01), air temperature (O: −0.691, p <0.01) and with north (O: 0.652, p <0.01), east (O: 0.588, p <0.01) coordinates in open areas. Precipitation correlation was also moderate with air temperature (O: 0.594, p <0.01; F: 0.591, p <0.01; B: 0.6, p <0.01) and north coordinate (O: −0.581, p <0.01; F: −0.510, p <0.01; B: −0.562, p <0.01) in all three sites types.

From all explanatory variables, the correlation between air temperature and north coordinate were strong and the absolute values exceeded 0.7 limits on all sites (O: −0.840, $p < 0.01$; F: −0.855, $p < 0.01$; B: −0.865, $p < 0.01$). This limit was surpassed also in all three sites between north coordinate and solar radiation (O: −0.878, $p < 0.01$; F: −0.791, $p < 0.01$; B: −0.789, $p < 0.01$). Due to these clear collinearities between explanatory variables, we decided to exclude solar radiation from all three models. Hence, we are able to reduce possible multicollinearity. However, we are keeping east and north coordinates in all three models because the spatial location may affect frost depth also through an effect mechanism other than temperature. Based on Holm's method, p-values of all highlighted correlations in matrixes were statistically significant (p-value <0.01) (Holm 1979).

6.1.2 Open area

Table 6 presents GAM model output for open areas. Here we can see that from all of our explanatory variables, NAO-index (*AVE.NAO*), site (*site*) variable, air temperature/snow interaction term (*AVE.Temp,AVE.MAX.Snow*) and spatial variables (*coord.east,coord.north*) were statistically significant. The diagnostic information about the fitting results of open area model is presented in Appendix 6a. R command *gam.check()* gave us four residual plots for each model. The degrees of freedom were adjusted on NAO index variable where we set the upper limit as *k=4* to smooth the estimate function line.

Table 6. GAM output for open areas model. Description in Table 4.

OPEN AREA					
AVE.MAX.Frostlayer					
	Parametric coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	28.59	2.324	12.3	<2e-16	***
Explanatory variables	edf	Ref.df	F	p-value	Signif.
s(AVE.Prec)	1.9	2.5	1.5	0.27729	
te(AVE.Temp,AVE.MAX.Snow)	17.2	22.1	14.8	< 2e-16	***
s(Freezing.Index)	1.0	1.0	0.3	0.57448	
s(AVE.NAO)	2.8	3.0	8.5	0.00017	***
s(coord.east,coord.north)	4.1	4.1	11.4	5.28e-9	***
s(season.id)	2.1	2.6	1.5	0.19281	
s(site)	29.0	32.0	22.7	< 2e-16	***
R-sq.(adj) = 0.893 Deviance explained = 90.4%					
GCV = 92.172 Scale est. = 83.415 n = 565					

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The GAM models estimated smooth functions are illustrated in Figures 3-6 where lines represent the effect of the explanatory variable to response (smooth line) and orange area shows a 95 % confidence interval. In Figure 4 we can see that on open areas, precipitation seems to have a decreasing effect to frost layer thickness but because of lack of observations, the confidence interval gets wider on both ends of the smooth line. Freezing index plot (Figure 4) does not show clear direction about its effect to response. The findings are in line with large p-values. Air temperatures and snow interaction plots are shown in Figure 5. Contour line distances are equidistant until snow cover becomes about 45 cm thick. Equal spacing is indicating that change of one degree on air temperature would have the same effect to the response regardless of the snow depth. The equidistant lines express that the temperature and snow effects to response are additive and linear. The vertical direction of contour lines implies that after snow layer exceeds 45-50 cm, snow does not appear to have impact on the response. However, lines horizontal distances are also getting wider which indicates that there would need to be a bigger change on air temperature for getting the same amount of effect to response when there is over 45-50 cm snow layer. As we can see from the second interaction plot, the number of observations is really low below −4 °C and less than 30 cm of snow which means we can’t be certain about actual interaction on this part.

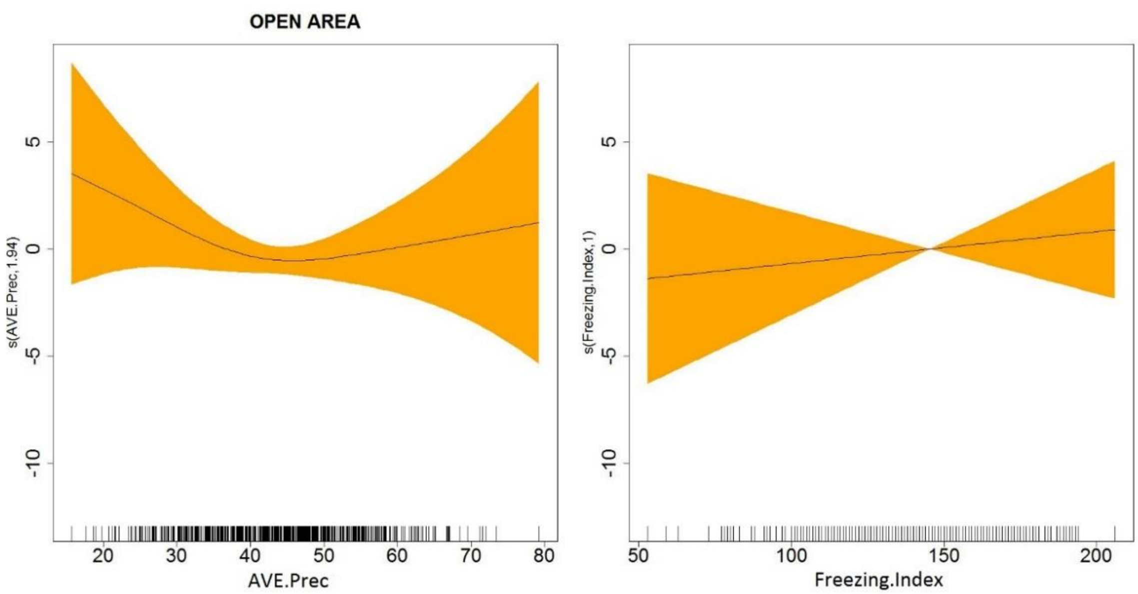


Figure 4. Estimated smooth functions for *AVE.Prec* and *Freezing.index* explanatory variables of open area model. The line is representing the smooth curve of the explanatory variable and orange area is represents the 95 % confidence interval.

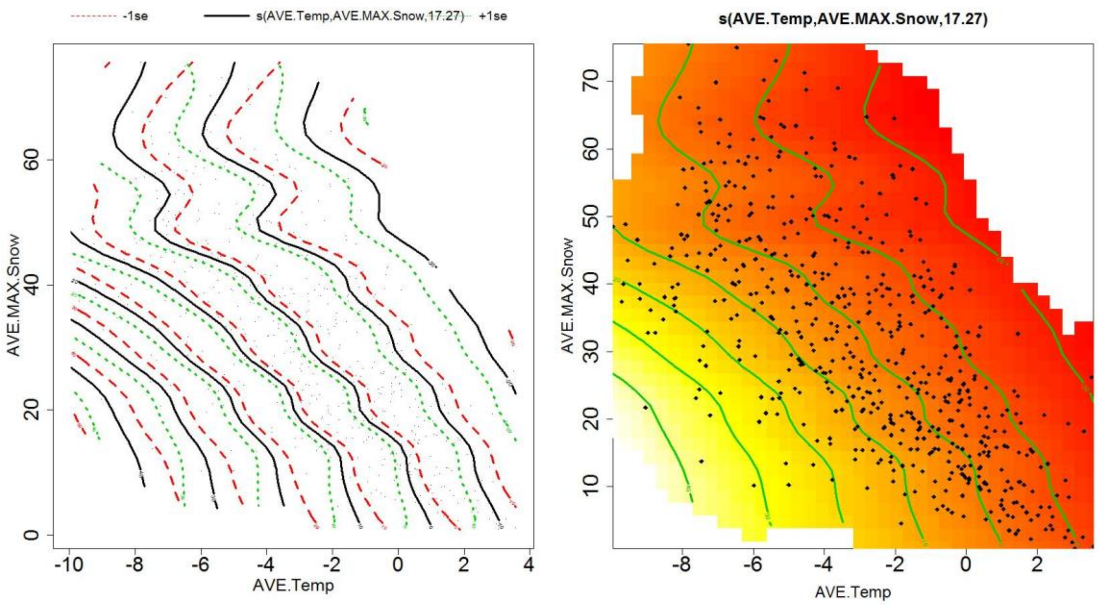


Figure 5. Open area: estimated smooth functions for tensor product smooth for mean maximum snow depth and mean temperature variables. The bold contours show the estimate of the smooth and dashed contours show the smooth plus standard error of the smooth and the dotted show the smooth less its standard error. The colour gradient representing response value: yellow = high, red = low. Black dots represent observations.

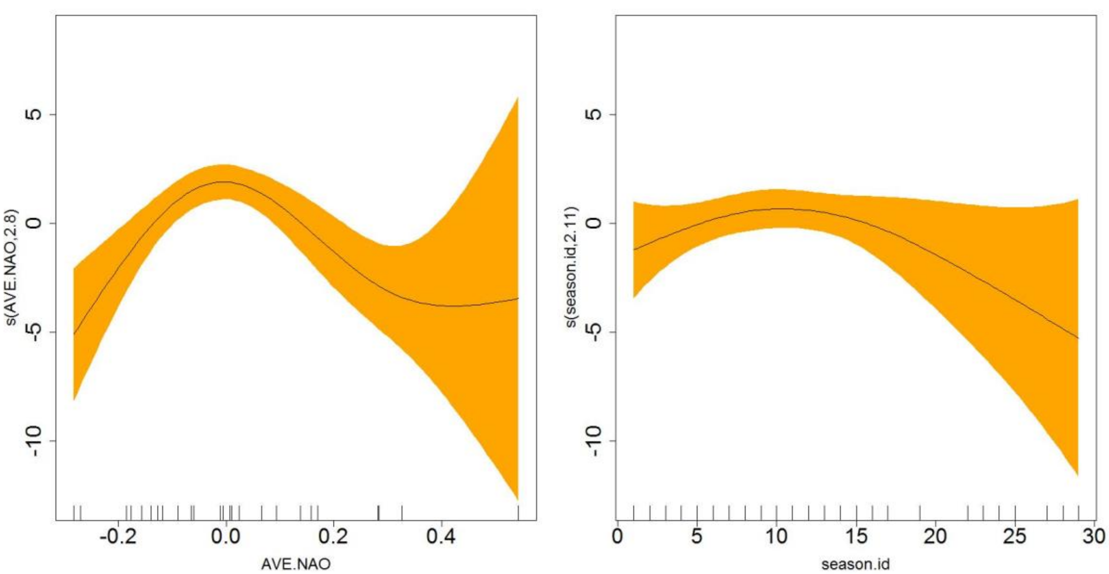


Figure 6. Estimated smooth functions for explanatory variables of open area model. Average NAO –index and year variable (*season.id*). The line is representing the smooth curve of the explanatory variable and orange area is represents the 95 % confidence interval.

NAO -index seems to have a positive effect on the response when the value is close to 0. Even though year variable (*season.id*) shows a declining trend on soil frost maximum thickness during the last 15 years the result is not statistically significant (Figure 6). Again we can see from x-axis that the number of observation is widening our confidence interval in NAO -index plot. Based on Figure 7, we can see that soil frost maximum thickness is increasing when we are moving toward the north as expected, but moving from east to west, there are little differences between eastern and western sites maximum.

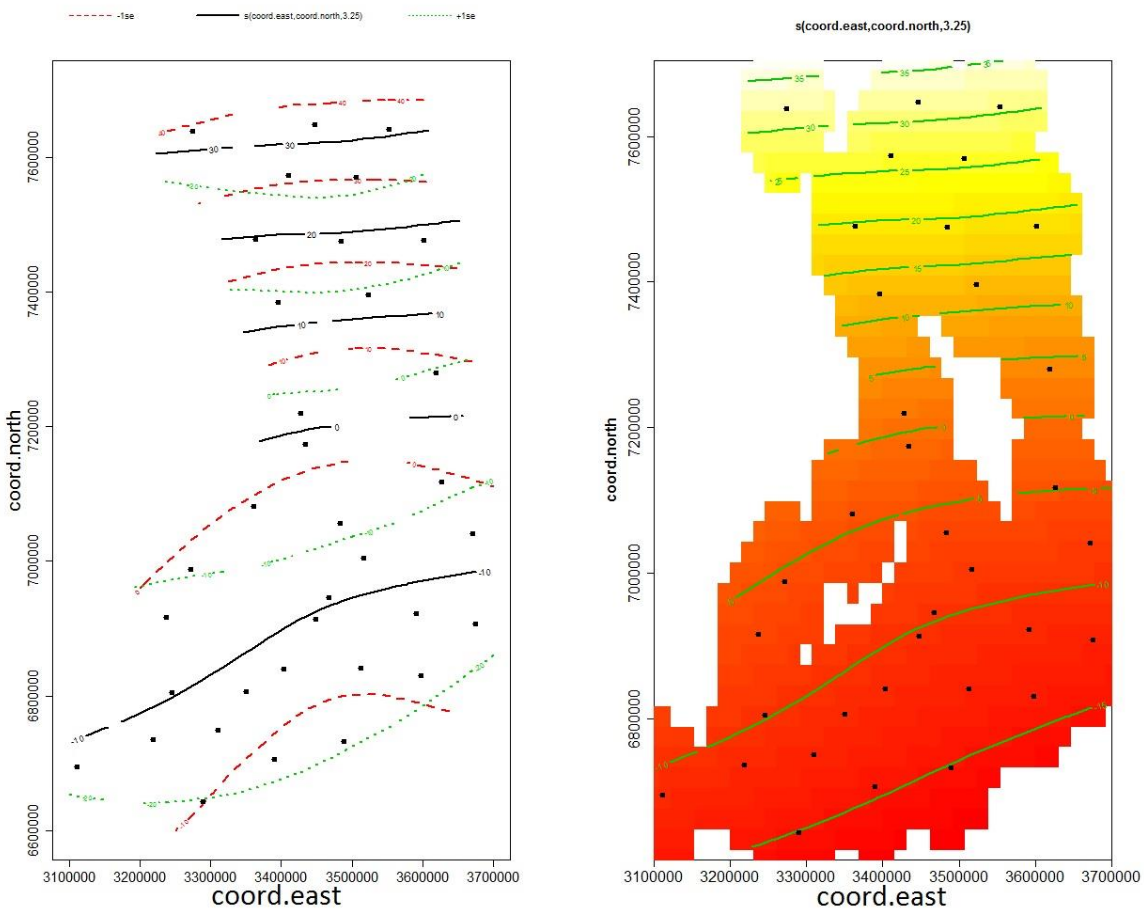


Figure 7. Estimated smooth functions of spatial variables for open area model. The colour gradient representing response value: yellow = high, red = low. Black dots represent sites spatial location.

Table 7. LME model output for open areas. *Snow.BIN.45*=snow depth categorized 0/1 variable based on 45 cm limit (0 = <45cm, 1 = >45cm). *Snow.BIN.45:AVE.Temp*= Interaction term where *Snow.BIN.45* represents soil frost when there is below 45cm of snow. *Snow.BIN.45:AVE.Snow*= interaction term where *Snow.BIN.45* represents soil frost when there is below 45cm of snow. *Prec.BIN.45* = categorized 0/1 variable based on 45 cm limit (0 = <45cm, 1 = >45cm). I(NAO^2)= quadratic term of NAO –index. Detailed variable description in Table 4.

OPEN AREA					
Linear mixed-effects model fit by REML					
AIC	BIC			LogLik	
4326.707	4399.976			-2146.354	
Random effects:	~1 | site				
	(Intercept)			Residual	
StdDev:	15.00345			9.391252	
Fixed effects:	Value	Std.Error	DF	t-value	p-value
(Intercept)	−248.75887	81.28908	518	−3.060176	0.0023
Snow.BIN.45	−18.50851	7.68021	518	−2.409898	0.0163
AVE.Temp	−4.51462	0.36714	518	−12.296752	<0.0001
AVE.MAX.Snow	−0.69490	0.06826	518	−10.180582	<0.0001
Prec.BIN.45	−18.30077	6.61468	518	−2.766691	0.0059
AVE.Prec	−0.21855	0.10829	518	−2.018128	0.0441
AVE.NAO	6.45428	3.15242	518	2.047403	0.0411
I(AVE.NAO^2)	−60.85917	13.47878	518	−4.515185	0.0000
Freezing.Index	0.03117	0.02670	518	1.167083	0.2437
season.id	−0.02360	0.08754	518	−0.269545	0.7876
coord.east	−0.00002	0.00002	32	−0.999440	0.3251
coord.north	0.00005	0.00001	32	5.679324	<0.0001
Snow.BIN.45:AVE.Temp	1.24504	0.54573	518	2.281417	0.0229
Snow.BIN.45:AVE.MAX.Snow	0.44581	0.14259	518	3.126478	0.0019
Prec.BIN.45:AVE.Prec	0.40871	0.14469	518	2.824708	0.0049

Signif. codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '+'

In Table 7 is presented LME model output for open areas. On LME model for open areas, *Freezing.Index*, *season.id* and *coord.east* variables were not statistically significant (p-values >0.05). Based on interaction plot the open areas snow layer limit seems to place around 40-45 cm. Soil frost has been categorized based snow depth where 0 represents observations <45 and 1 observations ≥45 centimetres. Based on *AVE.Temp* expected value, one-degree temperature increase on freezing season mean temperature leads to 4.51 cm decrease on soil frost layer when there is <45 cm of snow. When the snow layer thickness is ≥45cm the effect for average temperature can be calculated by summing the *Snow.BIN.45:AVE.Temp* value

(1.24) and *AVE.Temp* value (−4.51). This gives us the total value of −3.27 which means that one-degree increase on mean temperature when the snow layer thickness is ≥45 cm leads to around 3 cm decrease on soil frost. On average one-centimetre increase on freezing season average maximum snow depth (*AVE.MAX.Snow*) leads to 0.69 cm decrease on soil frost when snow cover is <45 cm. By following the same procedure with *Snow.BIN.45:AVE.MAX.Snow* and *AVE.MAX.Snow* we get the expected value 0.2 cm decrease on the response when freezing season average maximum snow cover is increasing by 1 cm in cases of ≤45 cm of snow cover. Based on the linear (*AVE.Prec*) term coefficients of precipitation, it has a negative effect on frost thickness maximum decreasing soil frost on average by −0.21 cm in the situation of <45 mm rainfall. The sum of *AVE.Prec* and *Prec.BIN.45:AVE.Prec* values represent the effect to response when there has been ≥45 millimetres rainfall and we increase the precipitation with one unit. In this situation, one unit increase would lead to a positive effect on soil frost increasing it on average by 0.19 when we sum up −0.218 and 0.408 centimetres. With quadratic term I(*AVE.NAO^2*) we can calculate the peak of the parabola which we could identify from Figure 5. The peak value can be calculated with the next equation

$$-b/(2a) \tag{13}$$

where *b* is representing original variable *AVE.NAO* and *a* is representing quadratic term I(*AVE.NAO^2*). We get as a peak value (− (6.45)/(2*−60.8) = 0.05) which correspondence well with our GAM plot. The average freezing index value in Finland is around 140 days for our study period. Now the results are indicating that one day increase on winter average freezing index leads to 0.03 cm increase on season's maximum soil frost depth on open areas however the result is not statistically significant. The *season.id* represents running year values 1981-2010 describing changes on response between this time periods. On open areas, there appears to be no significant linear trend. Based on east and north coordinates estimated regression coefficients frost layer thickness is decreasing (−0.00002) when moving one meter eastwards and increasing (0.00005) when moving northwards. In other words, this would mean 2 cm decrease when moving 100 km eastwards and 5 cm increase moving 100 km northwards on open areas. However, based on the p-value of results, only the north coordinate is statistically significant. The random effect term output values are indicating that the location is affecting to response.

6.1.2 Forest

The forest models output is presented in Table 8. The diagnostic information about the fitting results of forest model is presented in Appendix 6b. The degrees of freedom were adjusted from NAO index *(AVE.NAO)* and year *(season.id)* variables. Here we set as a limit $k=4$. Based on p-values, all of the variables were statistically significant (p-value <0.05) except freezing index. Plots of the estimated smooth functions are in Figure 8-11 and here we can see that precipitation has the same declining effect until season average rainfall (*AVE.Prec*) reach 45-50 mm (Figure 8). After this, the effect becomes more unclear due to lack of observations. The interaction term and its 2-dimensional plot are clearly different than on open areas. Lines in Figure 9 have almost equal spacing at below −2 °C and under 20 cm of snow layer, indicating that the change of one degree on air temperature would have the same effect to the response regardless snow depth. The lines direction towards upright in the plot would mean that at these temperatures and with this snow depth range, snow has overall increasing effect on response. When snow layer exceeds 20 cm, this impact is changing to negative. Overall lines are more vertically placed which would mean that snow does not have as big effect to response on forest as on open areas. However, we really cannot be certain about snow or air temperature interaction on under 20 cm snow cover and below −8 °C temperature due to lack of the observations. In Figure 10 year variable (*season.id*) is indicating decreasing trend on maximum depth of soil frost on forest sites.

Table 8. GAM output for forest model. Description in Table 4.

FOREST					
AVE.MAX.Frostlayer					
Parametric coefficients:					
	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	26.268	1.253	20.97	<2e-16	***
Explanatory variables	edf	Ref.df	F	p-value	Signif.
s(AVE.Prec)	4.1	5.2	9.4	7.22e-9	***
te(AVE.Temp,AVE.MAX.Snow)	15.1	17.5	12.2	< 2e-16	***
s(Freezing.Index)	1.0	1.0	0.1	0.8187	
s(AVE.NAO)	2.3	2.7	5.4	0.0033	**
s(coord.east,coord.north)	14.9	15.4	8.9	< 2e-16	***
s(season.id)	1.0	1.0	12.1	0.0005	***
s(site)	18.2	34.0	5.1	< 2e-16	***
R-sq.(adj) = 0.889 Deviance explained = 89.7%					
GCV = 86.239 Scale est. = 79.88 n = 783					

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

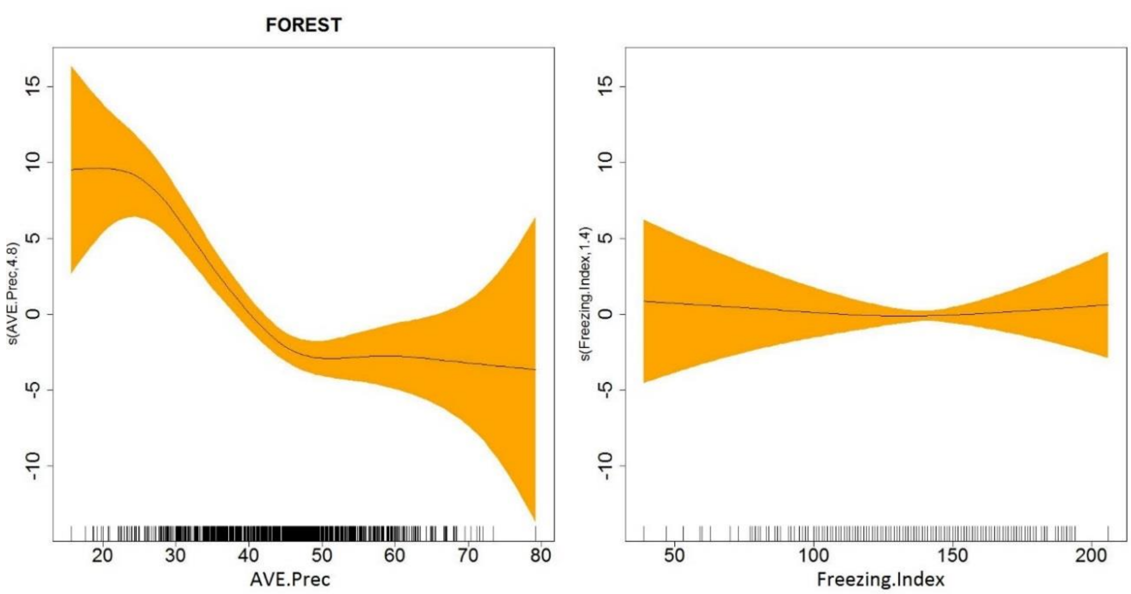


Figure 8. Estimated smooth functions for explanatory variables of forest model. The line is representing the smooth curve of the explanatory variable and orange area is represents the 95 % confidence interval.

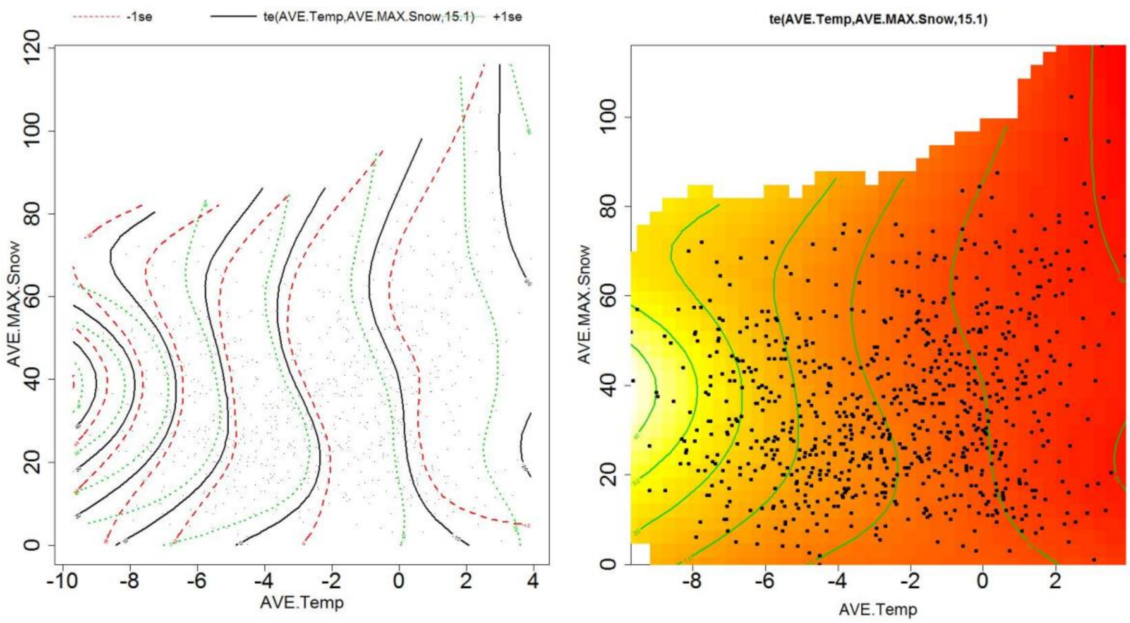


Figure 9. Forest: estimated smooth functions for tensor product smooth for mean maximum snow depth and mean temperature variables. The bold contours show the estimate of the smooth and dashed contours show the smooth plus standard error of the smooth and the dotted show the smooth less its standard error. The colour gradient representing response value: yellow = high, red = low. Black dots represent observations.

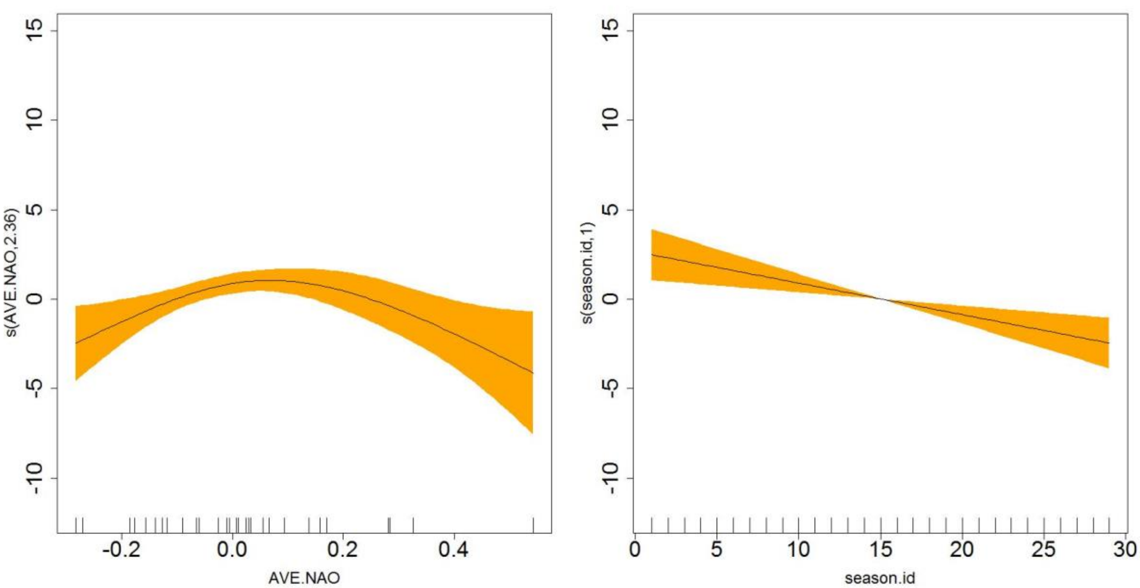


Figure 10. Estimated smooth functions for explanatory variables of forest area model. Average NAO –index and year variable (season.id).

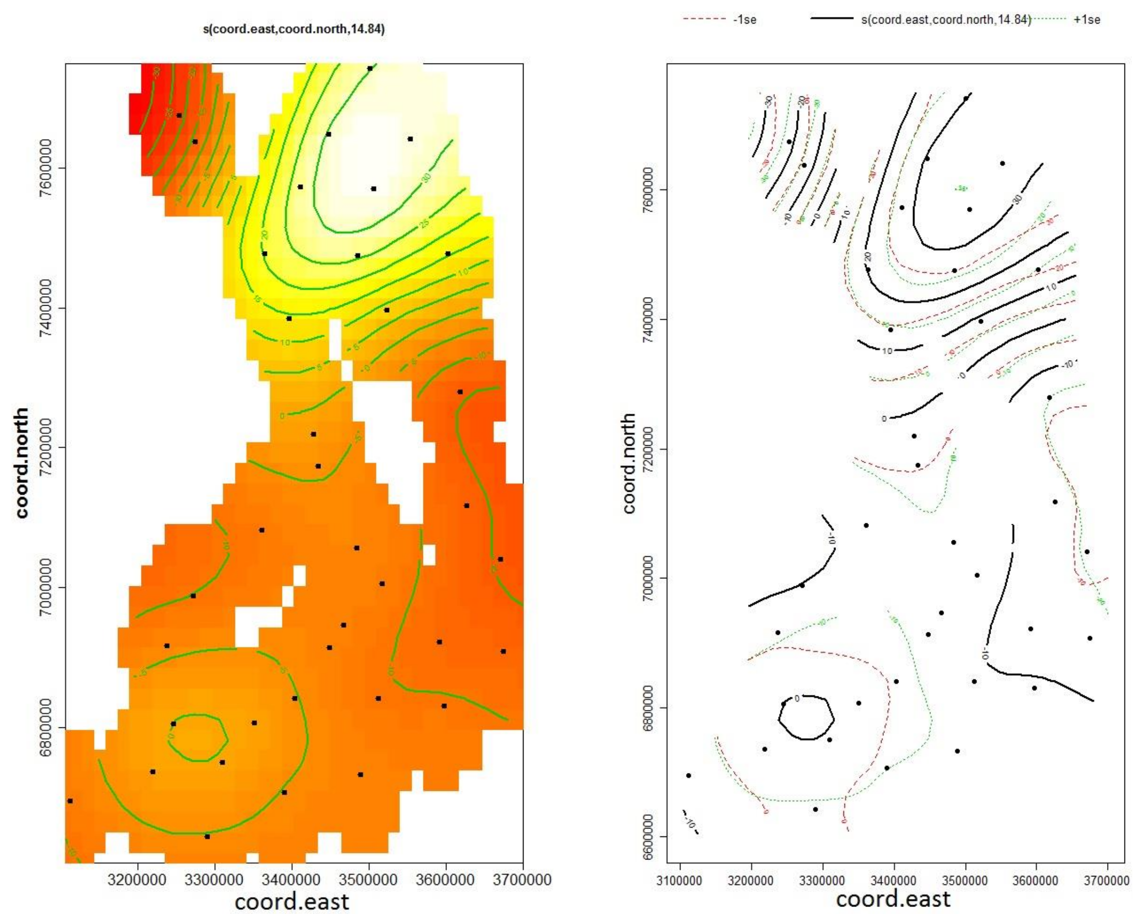


Figure 11. Plots of estimated smooth functions of spatial variables for forest model. The colour gradient representing response value: yellow = high, red = low. Black dots represent sites spatial location.

In Table 12 is presented LME output for frost tube sites located in the forest. Based on p-values, *Freezing.Index* and *coord.east* were not statistically significant (p-value>0.05). The selection of threshold was not as clear as on open. Because forest plot indicated in Figure 9 a positive effect to response during below 20 cm snow layer, we set this as a boundary for our LME model.

One degree increase on average temperature (*AVE.Temp*) on freezing season decreases soil frost maximum depth on average by 3.1 cm during <20 cm of snow cover. However *SNOW.BIN.20:AVE.Temp* interaction term showed that air temperature has an overall higher negative effect ((−1.10) + (−3.10) = −4.20) to response in conditions when there is ≥20 cm thick snow cover. Increasing average maximum snow depth

(*AVE.MAX.Snow*) by 1 cm leads to a minor decrease (0.08) on the response. In this time *SNOW.BIN.20:AVE.MAX.Snow* interaction term did not pass the p-value limit (0.05) and was excluded from the model. The linear precipitation term *AVE.Prec* decreases (0.46) soil frost depth when freezing season average amount of precipitation increases one unit (1 mm) on the measurement area and the amount of precipitation is <45 mm. Summing of linear term of precipitation and interaction term (0.55−0.46 = 0.09) is indicating again slight positive effect after seasons average rainfall surpasses 45 mm. We can now calculate the peak value of NAO-index in Figure 10 with original and quadratic term values (− (6.48)/(2*−29) = 0.11). The season average freezing index value does not have a significant effect on the response. Time variable *season.id* value is indicating decreasing trend by 0.17 cm per year of freezing season soil frost maximum thickness on forest sites. Spatial variables have similar effect to response as on open areas. East coordinate was again insignificant. North coordinate increases response by 3 cm when moving 100 km northwards (0.00003). The random effects estimated standard deviation indicated that site variable does have an impact to response.

Table 12. LME model output for forests. Snow.BIN.20=snow depth categorized 0/1 variable based on 20 cm limit (0 = <20cm, 1 = >20cm). Snow.BIN.20:AVE.Temp= Interaction term where Snow.BIN.20 represents soil frost when there is below 20cm of snow. Snow.BIN.20:AVE.Snow= Interaction term where Snow.BIN.20 represents soil frost when there is below 20cm of snow. I(NAO^2)= quadratic term of NAO –index. Detailed variable description in Table 4.

FOREST					
Linear mixed-effects model fit by REML					
AIC		BIC		LogLik	
5825.247		5899.359		-2896.623	
Random effects:	~1 | site				
	(Intercept)			Residual	
StdDev:	16.64613			9.114016	
Fixed effects:	Value	Std.Error	DF	t-value	p-value
(Intercept)	−106.57709	85.52913	726	−1.246091	0.2131
SNOW.BIN.20	0.08764	1.49557	726	0.058602	0.9533
AVE.Temp	−3.10817	0.44050	726	−7.055937	<0.0001
AVE.MAX.Snow	−0.08774	0.03745	726	−2.342580	0.0194
Prec.BIN.45	−25.39328	5.32111	726	−4.772175	<0.0001
AVE.Prec	−0.46569	0.08113	726	−5.739702	<0.0001
AVE.NAO	6.48447	2.66389	726	2.434211	0.0152
I(AVE.NAO^2)	−29.05188	7.74230	726	−3.752358	0.0002
Freezing.Index	−0.00234	0.02109	726	−0.111145	0.9115
season.id	−0.17532	0.05264	726	−3.330909	0.0009
coord.east	−0.00001	0.00002	33	−0.491610	0.6262
coord.north	0.00003	0.00001	33	2.955662	0.0057
SNOW.BIN.20:AVE.Temp	−1.10682	0.36011	726	−3.073541	0.0022
Prec.BIN.45:AVE.Prec	0.55818	0.11474	726	4.864686	<0.0001

Signif. codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '+'

6.2.3 Bog

In Table 13 we have model results for soil frost measurements on bogs. The diagnostic information about the fitting results of open area model is presented in Appendix 6c. The degrees of freedom where adjusted from NAO index *(AVE.NAO)* and year *(season.id)* variable. This included the least amount of statistically significant variables where only interaction term (*AVE.Temp,AVE.MAX.Snow*), NAO-index (*AVE.NAO*) and site (*site*) variables showed sufficient statistical evidence.

Table 13. GAM output for bog model. Description in Table 4.

BOG					
AVE.MAX.Frostlayer					
Parametric coefficients:					
	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	12.129	1.435	8.453	<2e-16	***
Explanatory variables	edf	Ref.df	F	p-value	Signif.
s(AVE.Prec)	1.9	2.4	1.5	0.24252	
te(AVE.Temp,AVE.MAX.Snow)	10.3	12.2	13.2	< 2e-16	***
s(Freezing.Index)	1.0	1.0	3.1	0.07914	.
s(AVE.NAO)	6.1	7.2	2.8	0.00575	**
s(coord.east,coord.north)	2.0	2.0	1.8	0.16801	
s(season.id)	1.0	1.0	1.7	0.18641	
s(site)	32.7	35.0	21.6	< 2e-16	***
R-sq.(adj) = 0.642 Deviance explained = 66.9%					
GCV = 46.148 Scale est. = 42.514 n = 714					

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The first estimated smooth function in Figure 12 for precipitation is a similar one as on open areas indicating declining effect until it reaches 50mm/average. Freezing index showed an increasing effect which was not statistically significant based on p-value (>0.05). The interaction plot is presented in Figure 13. The direction of the lines to the upper right is indicating positive effect for snow when snow layer is under 30 cm thick and when the air temperature is below 0 °C. Contour lines distances are equidistant until snow cover exceeds 30 cm, where smooth lines are towards upper left and meaning that the effect of snow turns to negative. Contour lines spacing are now widening meaning that during thicker snow cover there needs to be bigger change on air temperature to get the same impact to response. This is implying that the thin snow layer would have an increasing effect to soil frost thickness and after a certain threshold this effect would diminish and turn to negative. However, the amount of observation below 20 cm of snow is again relative scarce, which is most likely due to high amount of zero values.

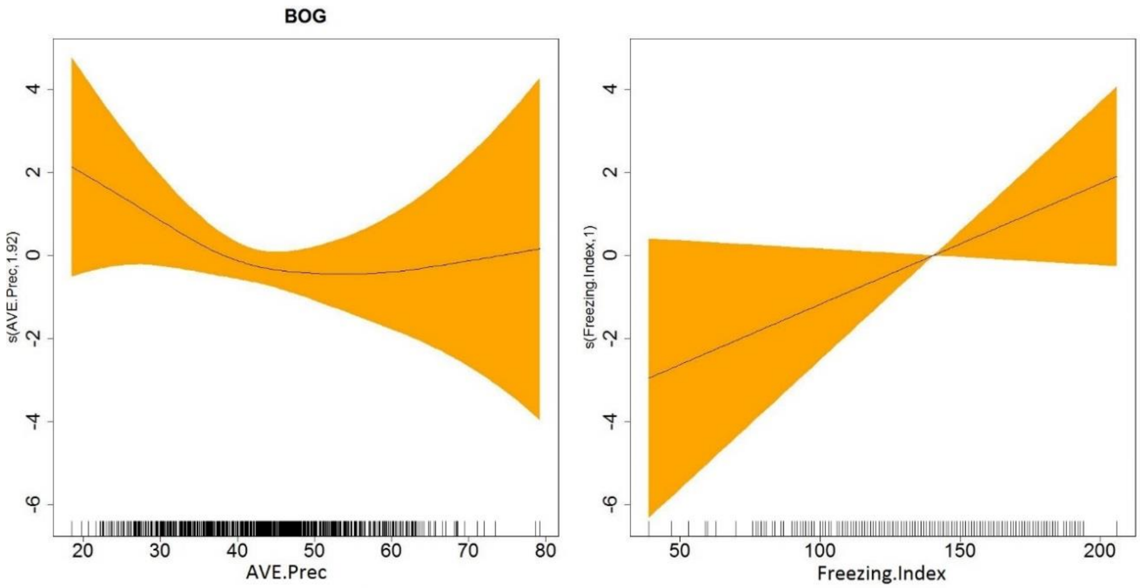


Figure 12. Plots of the estimated smooth functions for explanatory variables of bog model. The line is representing the smooth curve of the explanatory variable and orange area is represents the 95 % confidence interval.

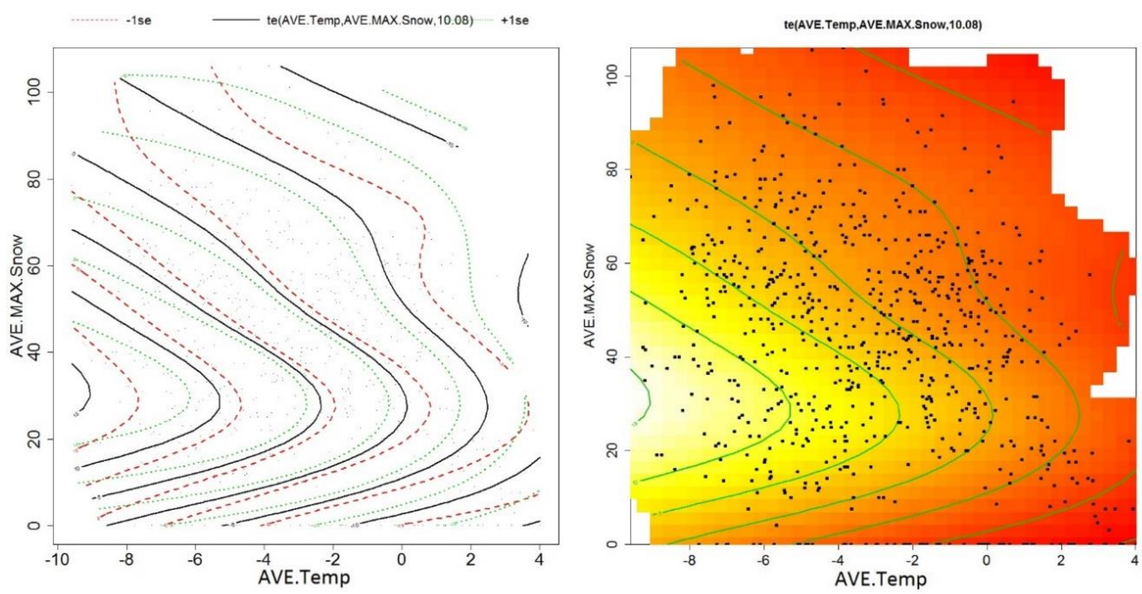


Figure 13. Plots of the estimated smooth functions for tensor product smooth for mean maximum snow depth and mean temperature variables. The bold contours show the estimate of the smooth and dashed contours show the smooth plus standard error of the smooth and the dotted show the smooth less its standard error. The colour gradient representing response value: yellow = high, red = low. Black dots represent observations.

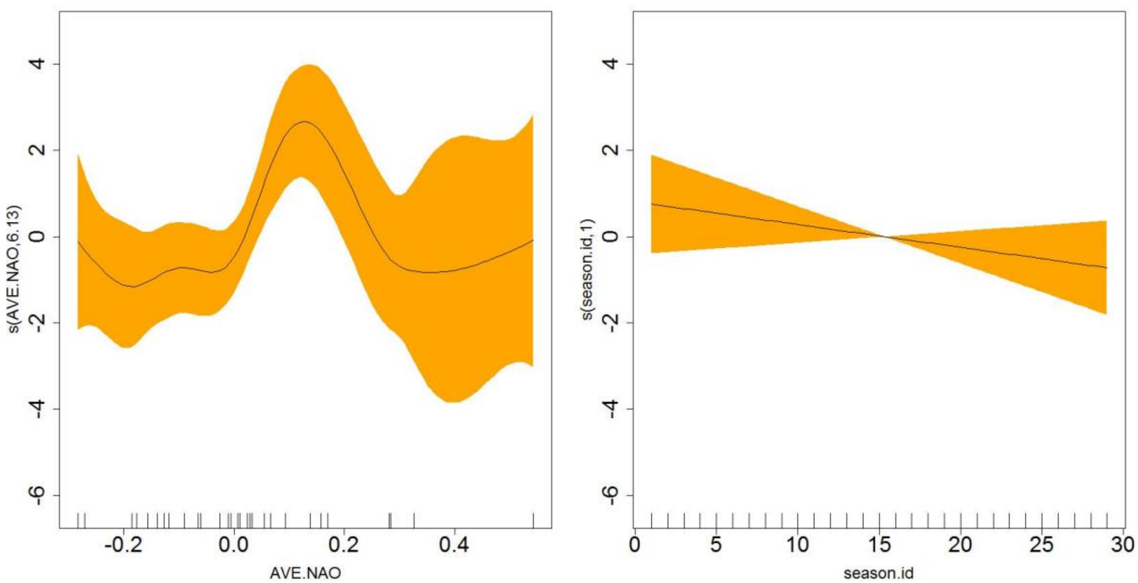


Figure 14. Plots of the estimated smooth functions for explanatory variables of bog model. The line is representing the smooth curve of the explanatory variable and orange area is represents the 95 % confidence interval.

The confidence intervals of NAO-index are again quite wide, but the overall form would express that during winter when the index value is between 0.0 and 0.2, the soil frost maximum thickness would be higher. Year variable (*season.id*) is showing that the soil frost maximum thickness during winter has declining trend 1981-2010, however, p-value suggests that the effect of time variable is not statistically significant (Figure 14). There is no significant spatial effect in frost depth in bog areas. Figure 15 is presenting spatial variables of the bog model and this indicates that frost is supposedly increasing southwards. However, the effect is weak and not statistically significant.

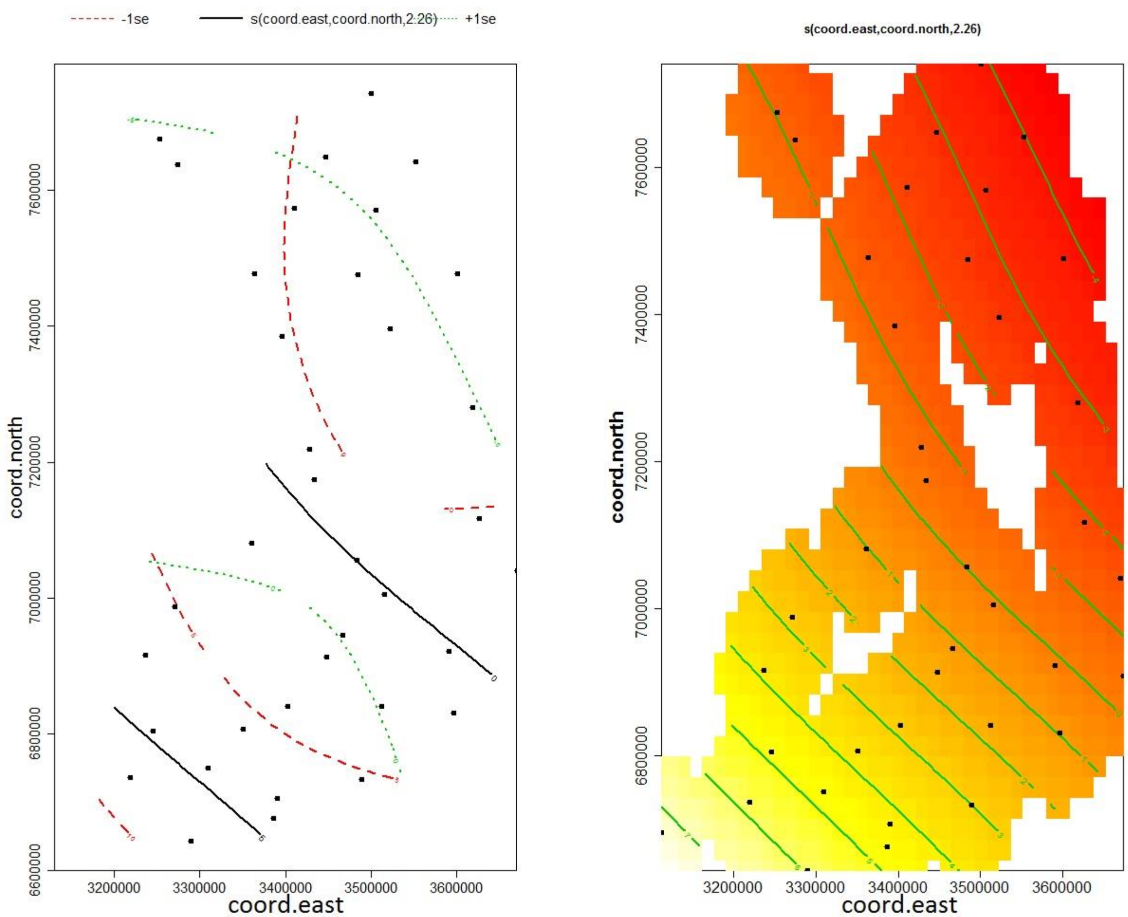


Figure 15. Plots of estimated smooth functions of spatial variables for bog model. The colour gradient representing response value: yellow = high, red = low. Black dots represent sites spatial location.

The LME model of bog frost tube is presented in Table 14. The bog interaction of GAM plot (Figure 14) showed a threshold when snow cover reaches around 30 cm depth. Below this, the snow layer effect seems to be positive to response and after 30 this turns to negative. From all variables, *Prec.BIN.45*, *AVE.NAO*, *Freezing.Index*, *season.id*, *coord.east* and *coord.north* were statistically insignificant.

The interaction term *SNOW.BIN.30:AVE.Temp* was excluded from the final model due to the high p-value. Mean air temperatures *AVE.Temp* effect to soil frost expected value is negative (−1.62) meaning that one-degree increase in freezing season mean air temperature decreases mean maximum soil frost on average of 1.6 cm. *AVE.MAX.Snow* expected value indicates that 1 cm increase on leads to a 0.4 cm increase on season's average soil frost maximum thickness during <30 cm snow cover. *SNOW.BIN30:AVE.MAX.Snow* interaction term sum with *AVE.MAX.Snow* ((−0.69) + 0.46 = −0.22) gives negative value meaning snow cover has a decreasing effect to frost layer during ≥30 cm snow layer. The linear precipitation term *AVE.Prec* indicates negative effect (−0.2) toward response when freezing season mean precipitation is increasing with one unit when mean rainfall is below 45 mm. In contrary to open and forest, if bog sites seasons average rainfall surpasses 45 mm, the effect of 1 mm increase on average to soil frost layer is negative (−0.03 cm). However this precipitation interaction terms p-value is close to the limit of statistical significance based on its p-value (0.0546) so its interpretation is questionable. Based on *season.id* variable there appears to be no significant trend. The spatial variables differ from two previous models indicating both a negative effect towards east and north based on regression coefficient estimates. However, the north gradient is milder than in other two areas and statistically less significant.

Table 14. LME model output for bogs. Snow.BIN.30=snow depth categorized 0/1 variable based on 30 cm limit (0 = <30cm, 1 = >30cm). Snow.BIN.30:AVE.Temp= Interaction term where Snow.BIN.30 represents soil frost when there is below 30cm of snow. Snow.BIN.30:AVE.Snow= Interaction term where Snow.BIN.30 represents soil frost when there is below 30cm of snow. Prec.BIN.45:AVE.Prec = interaction term where Prec.BIN.45 is categorized precipitation variable (0<45 and 1>45mm). I(NAO^2)= quadratic term of NAO –index. Detailed variable description in Table 4.

BOG					
Linear mixed-effects model fit by REML					
AIC	BIC		LogLik		
4907.179	4975.338		-2438.59		
Random effects: Formula	~1 | site				
	(Intercept)		Residual		
StdDev:	7.829578		6.72713		
Fixed effects: Formula	Value	Std.Error	DF	t-value	p-value
(Intercept)	107.0194	42.58599	660	2.513019	0.0122
SNOW.BIN.30	18.8989	1.88496	660	10.02615	<0.0001
AVE.Temp	−1.6218	0.25524	660	−6.35406	<0.0001
AVE.MAX.Snow	0.46874	0.06381	660	7.345517	<0.0001
Prec.BIN.45	−6.30232	4.18825	660	−1.504764	0.1329
AVE.Prec	−0.2089	0.06203	660	−3.367794	0.0008
AVE.NAO	3.26458	1.6945	660	1.926574	0.0545
Freezing.Index	0.01678	0.01663	660	1.009394	0.3132
season.id	−0.05976	0.03657	660	−1.634221	0.1027
coord.east	−0.00001	0.00001	35	−1.115878	0.2721
coord.north	−0.00001	0	35	−1.971796	0.0566
SNOW.BIN.30:AVE.MAX.Snow	−0.69358	0.07031	660	−9.86489	<0.0001
Prec.BIN.45:AVE.Prec	0.17292	0.08985	660	1.92467	0.0547

Signif. codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '+'

6.2 Northern Finland soil frost tubes in 1981-2010

6.2.1 Soil frost and snow cover maximum depth and frost season length

There can be seen clear differences between stations in northern Finland in Figures 16-18 where we have presented an average of soil frost layers maximum thickness in 30 years period in all three environmental types and in different sites. From Figure 16 we can see that in Inari Angel, Inari Nellim and Sodankylä Tähtelä sites soil frost means of yearly maximum is above overall average in the north on open areas but also variance is greater in these sites. On the other hand, Kemijärvi and Kittilä sites averages are clearly below overall mean and the variances are significantly smaller.

The differences between sites in forest are less distinct than in open areas even though there are some places which are clearly below average (Figure 17). Both sites in Inari, Sodankylä Tähtelä, Kittilä and Utsjoki's sites averages are over the mean. Like in the open area, Kemijärvi's site average is clearly below average also in forests. There is less variation in the lengths of the intequartile ranges (gray areas) on open areas except in Enontekiö's and Kemijärvi's site. The highest mean value of soil frost thickness on bog environment was in Inari A and Ylitornio. Lowest depth was in Enontekiö Iitto and Kemijärvi. From all of the three environment types, the mean of bogs soil frost layer thickness (15 cm) differed clearly from open areas (60 cm) and forests (60 cm). The lengths of interquartile ranges were quite small in all locations but the number of outliers was, on the other hand, higher than on two previous plots.

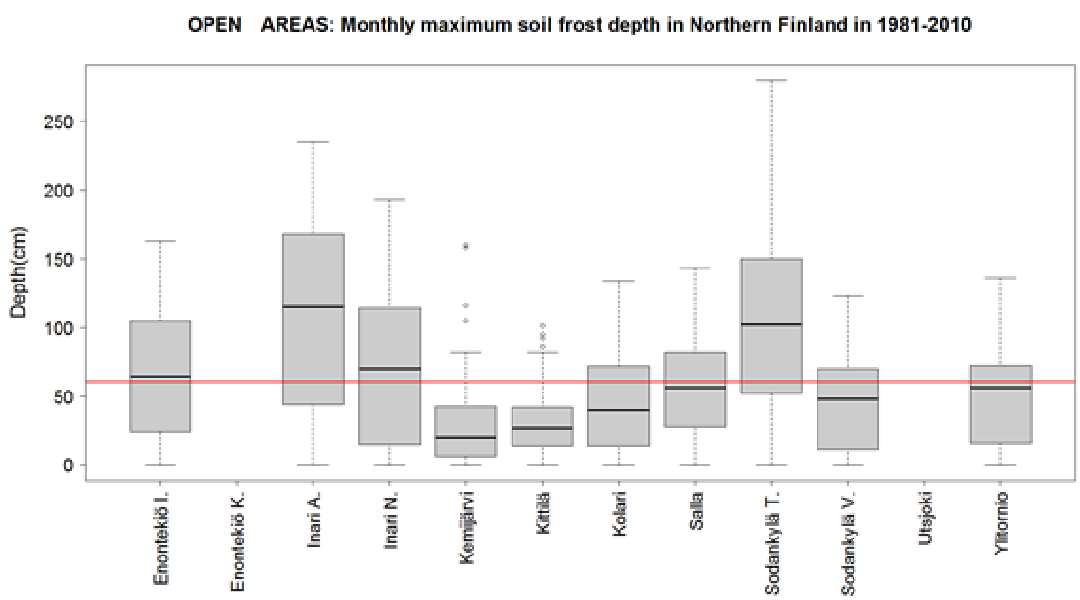


Figure 16. Monthly maximum soil frost depth on open area sites in northern Finland 1981-2010. Red line = Overall mean (60), Enontekiö Iitto (69.3), Inari Angel (108.4), Inari Nellim (68.5), Kemijärvi (28.8), Kittilä (29.5), Kolari (45.6), Salla (56.8), Sodankylä Tähtelä (101.3), Sodankylä Vuotso (44.3), Ylitornio (48.2).

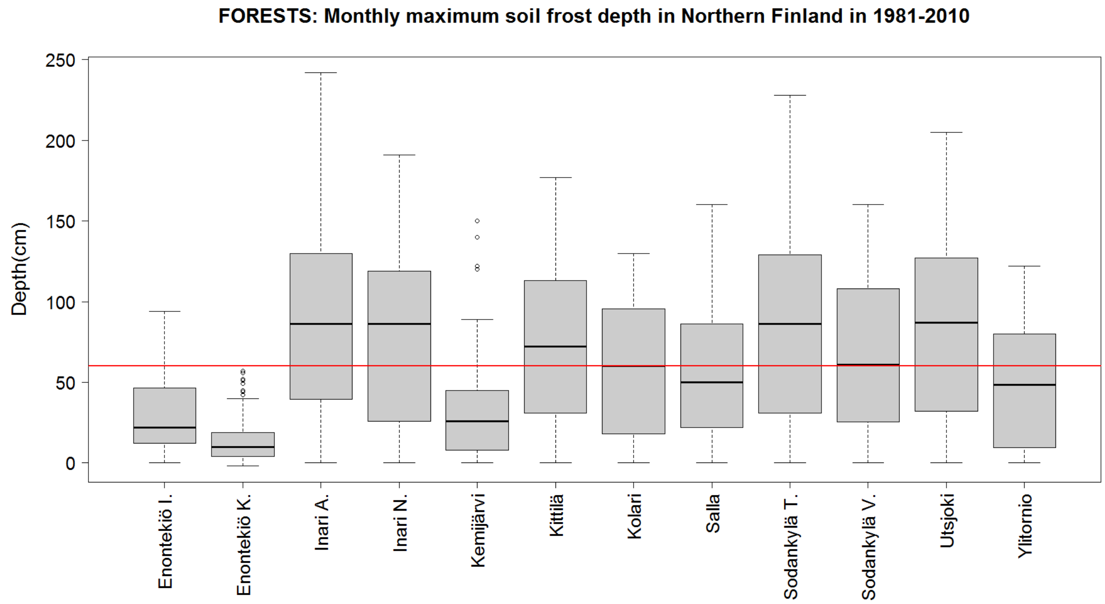


Figure 17. Monthly maximum soil frost depth on forest sites in northern Finland 1981-2010. Red line = Overall mean (60.6), Enontekiö Iitto (29.9), Enontekiö Kilpisjärvi (13.4), Inari Angel (89.1), Inari Nellim (78.3), Kemijärvi (29.3), Kittilä (71.9), Kolari (58), Salla (55.4), Sodankylä Tähtelä (83), Sodankylä Vuotso (66.9), Ylitornio (47.6).

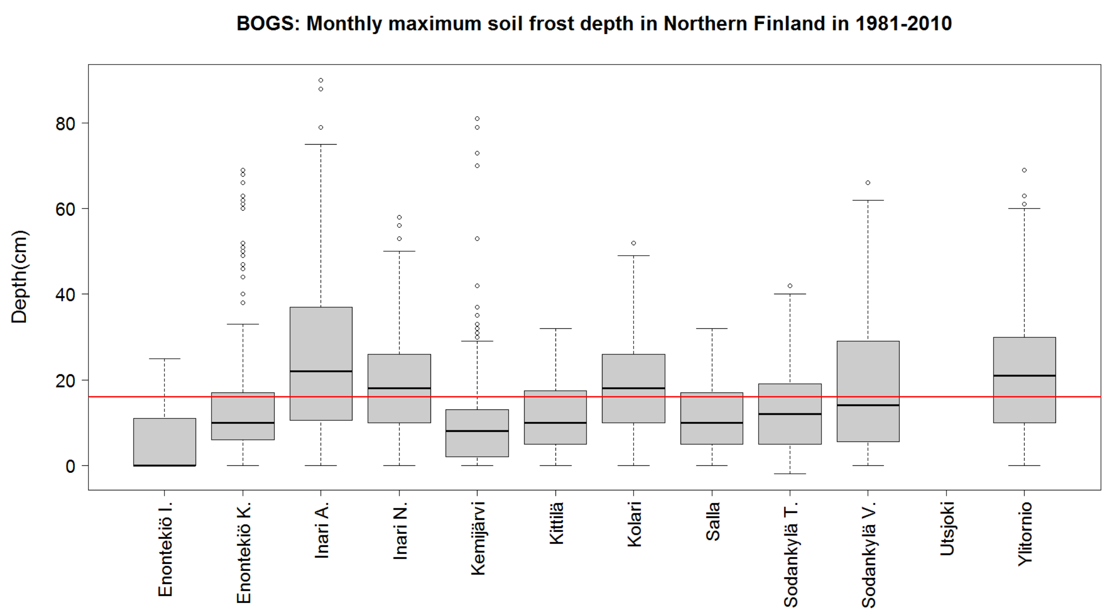


Figure 18. Monthly maximum soil frost depth on bog sites in northern Finland 1981-2010. Red line = Overall mean (16), Enontekiö Iitto (7.6), Enontekiö Kilpisjärvi (15.1), Inari Angeli (25.2), Inari Nellim (18.3), Kemijärvi (10.2), Kittilä (11), Kolari (18.3), Salla (55.4), Sodankylä Tähtelä (13.6), Sodankylä Vuotso (18.8), Ylitornio (21.3).

When we take all measurement sites with the same environmental type, we can illustrate whole northern Finland's frost tubes maximum frost depth in 30 year period in open areas, forests and bog sites. Figures 19-21 show yearly maximum depth of soil frost and snow cover on this time period with all three site type separately and also linear trend line of these where y-axis value 0 is representing the ground surface level. Soil frost depth values have been turned over to as negative values in Figures 19-21 where Min of Frost Depth represents actually maximum depth.

In first look on Figure 19, we can notice some similarities between yearly fluctuations on snow covers maximum and soil frost bottom layers maximum depth. For example, decrease on snow layers thickness after 1996 and 2000 led increasing soil frost bottom depth. Snow cover yearly maximum is indicating a decreasing trend (slope: −0.4 cm/year). Soil frost bottom edges clear positive trend (slope: 2.2 cm/year) this is indicating that soil frost overall layers thickness is decreasing in northern Finland on open areas.

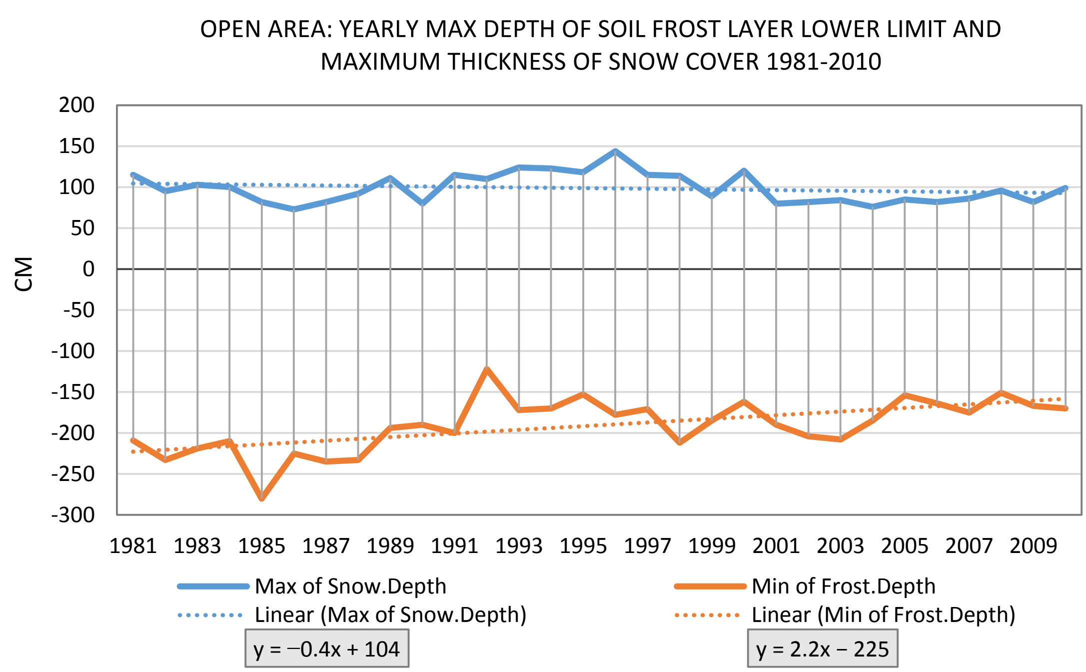


Figure 19. The average of yearly maximum depth of soil frost and maximum thickness of snow cover on open areas. The zero depth of y-axis represents ground surface. Min=Max Frost Depth.

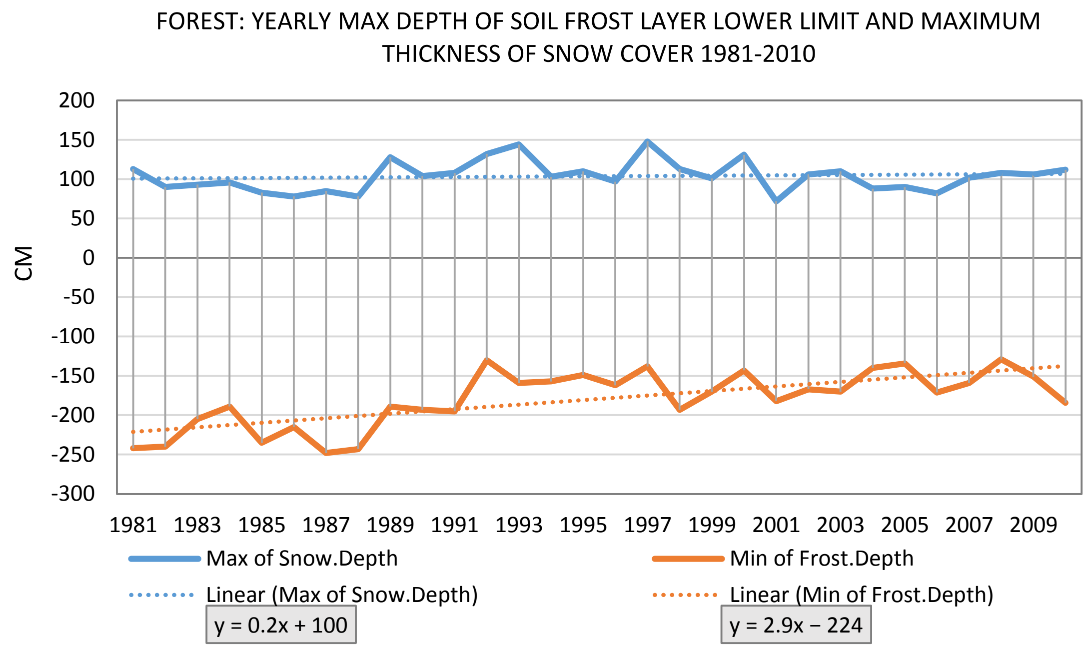


Figure 20. The average of yearly maximum depth of soil frost and maximum thickness of snow cover on forest sites. The zero depth of y-axis represents ground surface. Min=Max Frost Depth.

Forest frost tube sites yearly values and these linear trend lines are presented in Figure 20. From these, we can observe the same similarities on anomalies between frost layers depth and above snow cover thickness. There appear to be some differences in the yearly snow cover line compared to open areas. Decreasing snow depth often can be seen shallower soil frost but maybe not as clearly as open areas. Linear trend lines are suggesting a slight positive trend on open areas snow covers maximum depth. Likewise, to open areas, there is a clear positive and decreasing trend on soil frost bottom (2.9) boundary on forest sites.

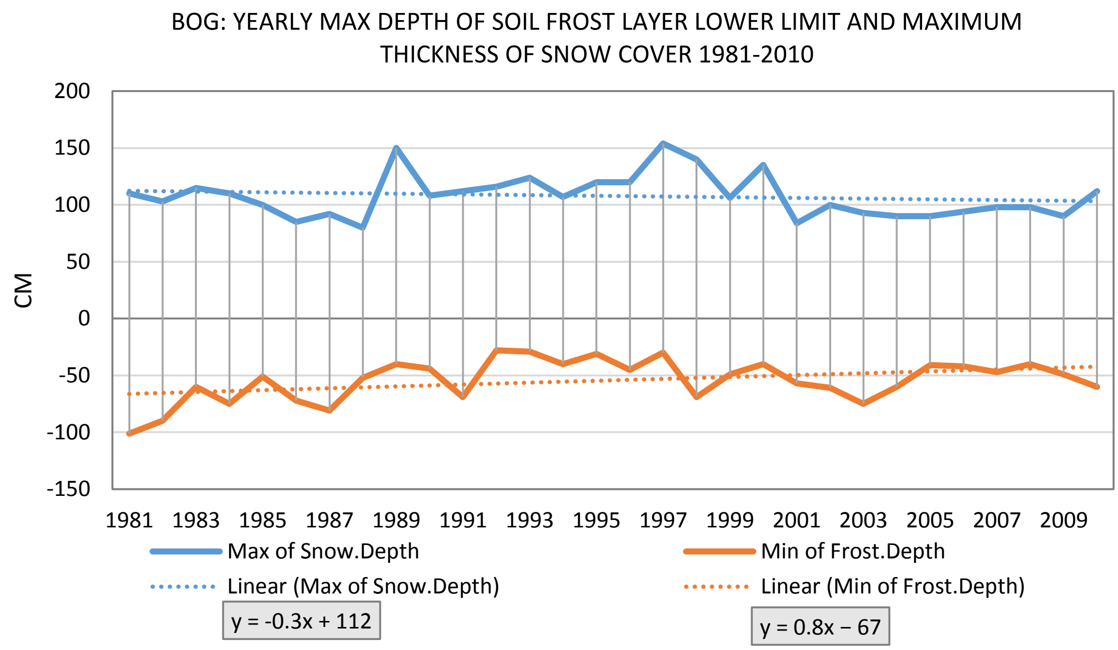


Figure 21. The average of yearly maximum depth of soil frost and maximum thickness of snow cover on bog sites. The zero depth of y-axis represents ground surface. Min=Max Frost Depth.

Frost tubes which were located on bog are presented in Figure 21. From this, we can detect that overall soil frost yearly minimum depth appears to be shallower on bogs compared to open areas or sites in the forest. Soil frost bottom line high and low points clearly follow and are affected by snow covers peaks similarly as in open areas and forest even though there are clear differences on soil frost layer boundaries fluctuations. Snow cover trend line is following the same negative pattern (−0.3) as on open areas. Soil frost bottom depth has decreased (0.8) like on both previous plots however not as clearly.

Frost season length in northern Finland and past changes during the 30 year period are presented in Figure 22. Different environmental types are presented separately. Red line represents the average of these three site types and linear trend line based on this averaged line. As we can observe from the plot, there are little differences between different environments. However, there is some clear fluctuation between frost seasons length and the overall trend of the number of days is slightly decreasing (−0.46). However, there is a significant amount of fluctuations during this 30 year period so stating the clear direction is fairly difficult.

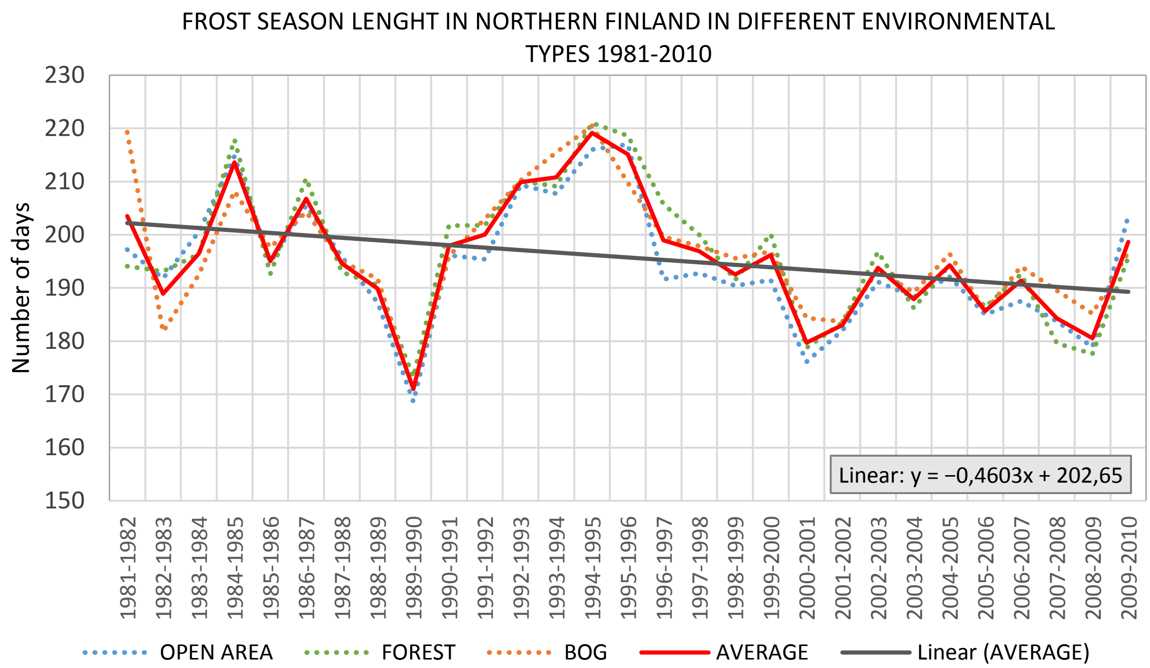


Figure 22. Winter frost season length in northern Finland between 1981 and 2010. Used period for each season September-April.

6.2.2 Mann-Kendall's trend and Sen's slope estimate test

We used Mann-Kendall's trend test and Sen's slope estimate to assess the significance of trends. In Figure 25 illustrates soil frost maximum depth change and its magnitude per year. The test statistics are presented in Table 16. Changes were mostly negative however statistically significant only on couple of sites. We were interested on sites which passed at least p-values 95 % ('*') confidence interval. There were three statistically significant trends in open areas, four on forests and two on bog sites and all of these were showing a decreasing trend. Especially Inari I.'s forest site indicated over 5 cm decrease on soil frost maximum depth and over 2 cm decrease on open site (Figures 23-25). Kittilä and Kolari sites on forest showed a decreasing trend as well as Kolari and Sodankylä V. bog site. There were also a couple of sites which showed a trend with 90 % confidence ('+') level and rest of weren't statistically significant (Table 16). The trends for soil frost maximum depth and frost season length for environmentally different frost tube sites are summed up in Table 15 and 16 and Figure 26. From these, we can observe that soil frost depth, as well as frost season length, have been decreasing on all three site types. The soil frost maximum depth has decreased on open areas approximately 2.12, on forest 2.75 and on bogs by 0.5 cm per year. Highest decrease was observed in the forest and lowest on bog sites. Soil frost depth changes were statistically significant (α=9v5) on open areas and forest. There were only small differences on frost season length on three types of sites and the diminishing was overall quite small on all three site types and this decrease was statistically significant only on a couple of places. From all of three site types, frost season decrease were statistically significant with 95 % confidence level (α=95) on open areas and with 90 % confidence level (α=90) on forests and bogs.

Table 15. Mann-Kendall's trend and Sen's slope tests results for the frost season length on open area, forest and bog sites.

	Mann-Kendall's trend test						Sen's slope estimate				
	Site	First year	Last Year	n	Test-Z	Signific.	Q	Qmin99	Qmax99	Qmin95	Qmax95
Frost Season Lenght	OPEN	1981	2009	29	-2.49	*	−0.51	−1.08	0.05	−0.88	−0.17
	FOREST	1981	2009	29	-1.93	+	−0.57	−1.32	0.14	−1.14	0.01
	BOG	1981	2009	29	-1.71	+	−0.48	−1.16	0.17	−0.92	0.07
p-value Signif. codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '+'											

Table 16. Mann-Kendall's trend and Sen's slope tests results for the maximum depth of soil frost in three different environment. Sites with less than 9 observations were left out (empty cells). NULL = missing frost tube. Colour gradient: white-orange increasing-decreasing trend.

				Mann-Kendall		Sen's slope estimate					
Site	First year	Last Year	n	Test-Z	Signific.	Q	Qmin99	Qmax99	Qmin95	Qmax95	
ENONTEKIÖ I.	1981	2010	20	-2.83	**	-2.02	−3.27	1.00	−2.82	0.00	OPEN AREA
ENONTEKIÖ K.	1981	2010	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
INARI A.	1981	2010	30	-4.27	***	-2.64	−3.83	−1.43	−3.53	−1.76	
INARI N.	1981	2010	30	-0.32		-0.13	−2	1.93	−1.35	1.44	
KEMIJÄRVI	1981	2010	30	-2.04	*	-1.00	−2.52	0.49	−2.05	0.07	
KITTILÄ	1981	2010	30	-1.21		-0.58	−1.875	0.52	−1.57	0.25	
KOLARI	1981	2010	30	-0.73		-0.33	−2	1.30	−1.67	0.85	
SALLA	1981	2010	30	1.09		0.94	−1.20	3.24	−0.64	2.68	
SODANKYLÄ T.	1981	2010	30	0.07		0.04	−1.88	2.00	−1.36	1.44	
SODANKYLÄ V.	1981	2010	23	-1.40		-1.36	−3.39	0.71	−2.92	0.10	
UTSJOKI	1981	2010	2			-2.00					
YLITORNIO	1981	2010	30	-0.91		-0.21	−1.26	0.40	−1	0.24	
Total	1981	2010	30	-3.52	***	-2.12	−3.63	−0.83	−3.08	−1.21	
ENONTEKIÖ I.	1981	2010	20	-0.78		-0.68	−2.41	1.25	−1.95	0.85	FOREST
ENONTEKIÖ K.	1981	2010	23	-1.67	+	-0.79	−1.80	1.09	−1.38	0.59	
INARI A.	1981	2010	30	-4.94	***	-5.13	−6.78	−3.30	−6.40	−3.86	
INARI N.	1981	2010	30	-0.70		-0.42	−1.85	1.18	−1.5	0.85	
KEMIJÄRVI	1981	2010	30	0.00		0.00	−1.29	0.87	−1	0.61	
KITTILÄ	1981	2010	30	-2.39	*	-1.63	−3.38	−0.016	−2.86	−0.30	
KOLARI	1981	2010	29	-3.61	***	-1.50	−2.69	−0.25	−2.42	−0.47	
SALLA	1981	2010	29	-1.43		-0.89	−2.69	0.97	−2.13	0.40	
SODANKYLÄ T.	1981	2010	30	-0.79		-0.50	−2.70	1.15	−2.15	0.73	
SODANKYLÄ V.	1981	2010	23	-0.34		-0.47	−3.20	2.44	−2.18	1.80	
UTSJOKI	1981	2010	30	-0.14		-0.09	−2	1.99	−1.68	1.42	
YLITORNIO	1981	2010	30	-2.13	*	-0.67	−1.46	0.13	−1.21	−0.10	
Total	1981	2010	30	-3.53	***	-2.75	−4.36	−1.19	−4	−1.62	
ENONTEKIÖ I.	1981	2010	6								BOG
ENONTEKIÖ K.	1981	2010	24	0.13		0.07	−1.33	1.99	−0.81	1.54	
INARI A.	1981	2010	30	-0.28		-0.28	−1.52	1.29	−1.18	1.00	
INARI N.	1981	2010	30	-1.40		-0.19	−0.71	0.19	−0.52	0.05	
KEMIJÄRVI	1981	2010	30	-1.34		-0.30	−0.88	0.33	−0.60	0.22	
KITTILÄ	1981	2010	30	-0.52		-0.14	−0.56	0.39	−0.45	0.25	
KOLARI	1981	2010	30	-2.07	*	-0.42	−0.90	0.25	−0.80	0.05	
SALLA	1981	2010	30	-0.63		-0.13	−0.42	0.36	−0.33	0.25	
SODANKYLÄ T.	1981	2010	30	0.56		0.13	−0.26	0.89	−0.14	0.77	
SODANKYLÄ V.	1981	2010	23	-3.18	**	-3.29	−3.26	0.00	−2.86	−0.25	
UTSJOKI	1981	2010	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
YLITORNIO	1981	2010	30	0.54		0.19	−0.67	0.77	−0.5	0.60	
Total	1981	2010	30	-1.34		-0.50	−1.63	0.56	−1.33	0.20	
p-value Signif. codes: 0.001'***', 0.01 '**', 0.05'*', 0.1 '+'											

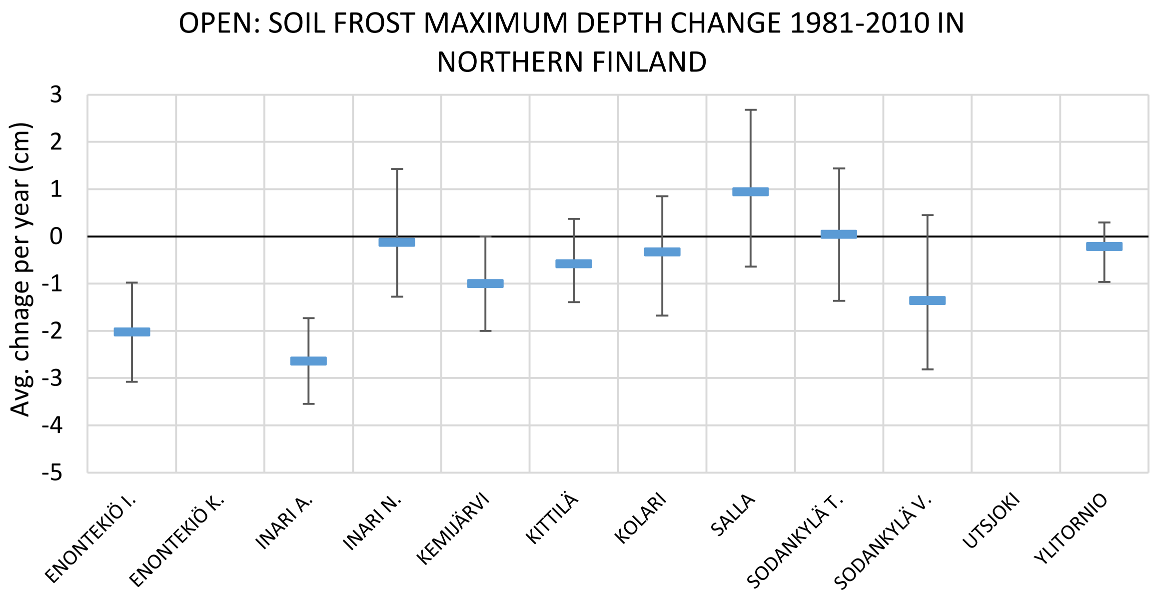


Figure 23. Soil frost maximum depths change in open areas in northern Finland. Y-axis: Sen's test (Q) slope estimator for the true slope of linear trend change per unit time period (year). Error bars: 95 % confidence intervals of Q. The blue horizontal line is representing the average value.

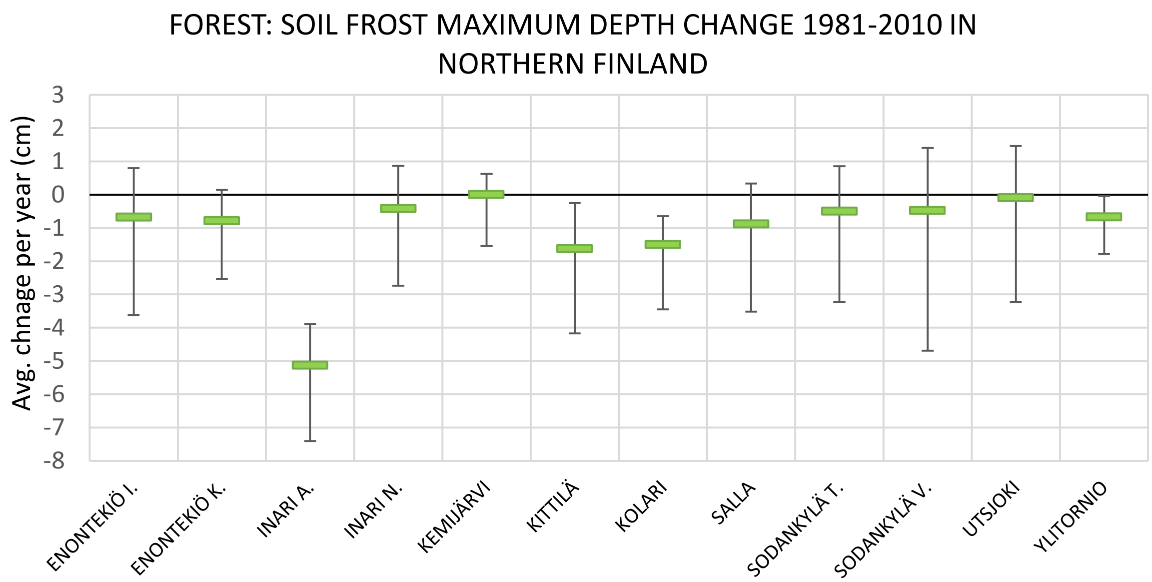


Figure 24. Soil frost maximum depths change in forest in northern Finland. Y-axis: Sen's test (Q) slope estimator for the true slope of linear trend change per unit time period (year). Error bars: 95 % confidence intervals of Q. The green horizontal line is representing the average value.

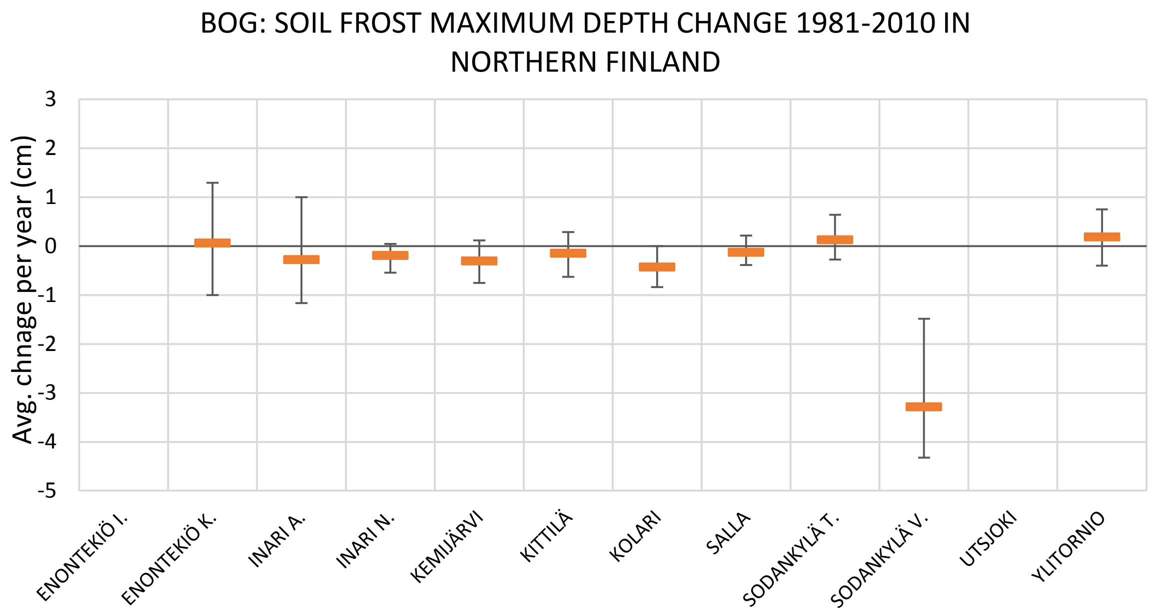


Figure 25. Soil frost maximum depths change in bogs in northern Finland. Y-axis: Sen's test (Q) slope estimator for the true slope of linear trend change per unit time period (year). Error bars: 95 % confidence intervals of Q. The orange horizontal line is representing the average value.

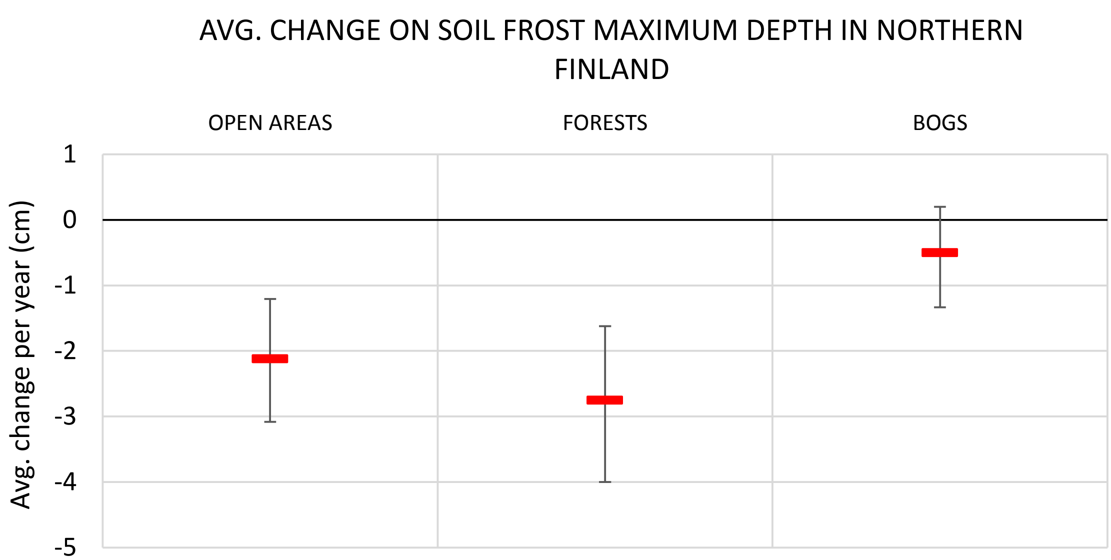


Figure 26. Total change on soil frost maximum depth in different environmental sites in northern Finland. Y-axis: Sen's test slope estimator (Q) for the true slope of linear trend change per unit time period (year). Error bars: 95 % confidence intervals of Q. The red horizontal line is representing the average value.

Table 17. Mann-Kendall's trend and Sen's slope tests results for the maximum depth of soil frost in different months. Sites with less than 9 observations were left out (empty cells). Colour gradient: white-orange increasing-decreasing trend.

			Mann-Kendall			Sen's slope estimate					
Month	First year	Last Year	n	Test-Z	Signific.	Q	Qmin99	Qmax99	Qmin95	Qmax95	
January	1981	2010	30	-3.68	***	-2.00	−2.95	−0.80	−2.70	−1.15	OPEN AREA
February	1981	2010	30	-3.36	***	-2.11	−3.22	−0.58	−2.91	−1.00	
March	1981	2010	30	-2.93	**	-1.83	−3.40	−0.47	−2.92	−0.80	
April	1981	2010	30	-3.75	***	-2.17	−3.71	−0.94	−3.15	−1.25	
May	1981	2010	30	-3.27	**	-2.22	−3.50	−0.69	−3.10	−1.07	
June	1981	2010	30	-1.13		-0.87	−2.68	1.12	−2.23	0.58	
July	1981	2010	21	-1.56		0.00	−6.06	0.00	−4.21	0.00	
August	1981	2010	6	NULL		NULL					
September	1981	2010	8	NULL		NULL					
October	1981	2010	30	-0.27		-0.05	−0.73	0.60	−0.51	0.40	
November	1981	2010	30	-0.77		-0.43	−2.23	1.03	−1.65	0.69	
December	1981	2010	30	-2.84	**	-1.48	−2.53	−0.14	−2.2	−0.43	
January	1981	2010	30	-3.93	***	-2.07	−3.08	−0.82	−2.78	−1.06	FOREST
February	1981	2010	30	-3.14	**	-1.94	−3.35	−0.41	−3.00	−0.80	
March	1981	2010	30	-3.75	***	-2.62	−4.08	−0.97	−3.77	−1.28	
April	1981	2010	30	-3.68	***	-2.76	−4.26	−1.23	−3.94	−1.71	
May	1981	2010	30	-4.16	***	-3.23	−4.73	−1.67	−4.46	−2.00	
June	1981	2010	30	-2.64	**	-2.11	−4.58	−0.11	−3.91	−0.71	
July	1981	2010	22	-0.69		-0.50	−7.45	4.45	−5.69	2.28	
August	1981	2010	6	NULL		NULL					
September	1981	2010	9	NULL		NULL					
October	1981	2010	30	-1.00		-0.23	−0.73	0.36	−0.54	0.20	
November	1981	2010	30	-1.46		-0.55	−1.46	0.50	−1.21	0.27	
December	1981	2010	30	-3.81	***	-1.82	−2.60	−0.75	−2.38	−1.00	
January	1981	2010	30	-1.79	+	-0.50	−1.22	0.24	−1.00	0.00	BOG
February	1981	2010	30	-1.77	+	-0.50	−1.44	0.25	−1.20	0.02	
March	1981	2010	30	-1.32		-0.45	−1.36	0.48	−1.05	0.24	
April	1981	2010	30	-1.14		-0.33	−1.33	0.39	−1.12	0.18	
May	1981	2010	30	-1.70	+	-0.70	−1.78	0.29	−1.47	0.09	
June	1981	2010	30	0.57		0.14	−0.67	0.86	−0.44	0.64	
July	1981	2010	19	-0.74		0.00	−2.63	1.25	−2.00	0.00	
August	1981	2010	4	NULL		NULL					
September	1981	2010	6	NULL		NULL					
October	1981	2010	30	1.16		0.20	−0.25	0.63	−0.11	0.55	
November	1981	2010	30	0.93		0.17	−0.35	0.53	−0.20	0.42	
December	1981	2010	30	0.13		0.00	−0.50	0.50	−0.33	0.37	

p-value Signif. codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '+'

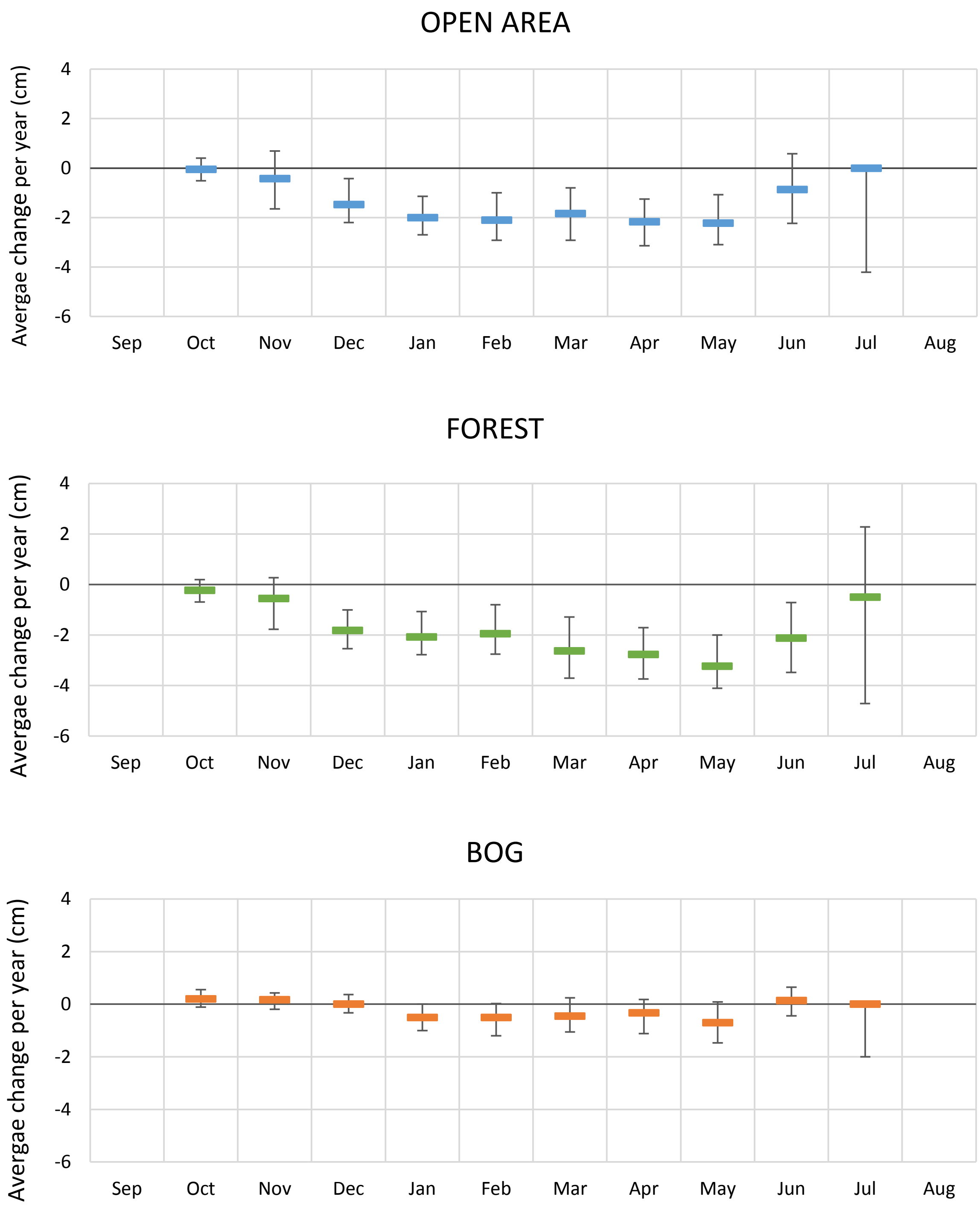


Figure 27. The average change in max of soil frost depth per year in different months in northern Finland.1981-2010 Y-axis: Sen's test (Q) slope estimator for the true slope of linear trend change per unit time period (year). X-axis from September to August. Error bars: 95 % confidence intervals of Q

We also analysed possible soil frost maximum depth changes in different months separately for three types of environments and these results are presented in Figure 27 and in Table 17. With a 99 % confidence interval there occurred statistically significant decreasing trend on months between December and May on open areas soil frost maximum depth. Sen's test Q-value indicates on average almost 2 cm decrease on soil frost max depth per year. Based a 99% confidence interval, the decrease occurs on forest sites between December and June. This decrease is around 2 cm per year based on Q-value. Bog sites did not show any over α=90 % decrease. January, February, and May were only cases with 90 % confidence interval and the estimated decrease per year was 1.5 cm. Due to lack of observations in September and August, we excluded these months in analysis on all three site types.

6.3 The applicability of Earth System Data Record for Land Surface Freeze/Thaw State and ERA Interim reanalysis of soil temperature

The results of correspondence between FT-ESDR and in-situ observations are presented in Table 18 and Appendix 7. From here we can observe that during 0 cm of soil frost, the percentage of FT-ESDR's incorrect predictions varied 17-83 % in different classes (Appendix 7). When overlaying snow cover is getting thicker, the accuracy of satellite evaluation is getting better especially with 0=frozen values. When there is soil frost but no snow cover present in the ground, the error rates of observations are as high as 60-76 %. However when increasing the soil frost depth and during shallow snow cover (1-5) the validity of observations isn't getting better as quickly as with snow cover. During 1-5 cm thick soil frost the error rate varies within 49-71 % when there is 1-5 cm thick snow cover is present. When snow cover is increasing the predictions correspondence with in-situ observations is clearly getting more accurate. For instance, regardless of snow cover and during 6-10 cm thick frozen ground layer, error rates of satellite predictions with frozen ground observations are varying within 17-65 %, on 11-20 cm 13-71 % and during over 20 cm 5-60 %. On the other hand, it was noticeable that when there was no frozen ground present the accurate thaw state predictions decreased with increasing snow cover. The model had overall a good ability to observe the thawing ground state without or during thin snow cover. However when snow layer became 6-10 cm thick only 43 % of predictions were accurate. The inaccuracy of FT-ESDR satellite predictions steadily decreased in our all five soil frost class and the mean varied within 27-41 %. When we look at the cases where frozen ground is present, the accuracy of predictions varied on snow classes: 0; 24-40 % 1-5; 49-71 %, 6-10; 61-75 %, 11-20; 70-84 % and >20; 83-95 %.

ECMWF soil temperature predictions presented against in-situ observations in Appendix 8. These predictions are describing the temperature of the ground at first 7 cm from the soil surface. In the first section has been presented cases where the ground is on thaw state. Again we can notice same increasing error rate with increasing snow cover where the percent vary between 25 and 85 %. When we are looking at values distribution during frozen ground we can again notice how correspondence is varying clearly with the snow covers thickness. When there is no snow cover present, the error rate of predictions with in-situ observations are again clearly high and percent of accurate observations are not increasing as steadily as with snow cover. When the frozen ground layer is getting thicker but snow cover is staying at zero, the error rate of observation percentages is varying 43-68 %. Correspondences are steadily increasing in all soil frost depth classes (6-10, 11-20, >20) when snow cover is getting thicker. In soil frost class 6-10 error rate percentages are varying within 8-68 %, in 11-20 class 8-43 % and >20 class 4-50 % regardless of snow cover. The percentages are again indicating that when there is none or shallow snow and soil frost depths, predicting soil frost presence with soil temperature is quite inaccurate compared to a situation where there is more snow and several centimetres thick layer of soil frost in the ground. In Figure 28 is presented the mean error rates for each soil frost class.

In Appendix 9 we lastly described the correspondence of FT-ESDR and ERA-Interim data and in Table 19 is presented the percentage of disagreeing observations. As we can see the lowest matching percentage is during shallow snow cover and soil frost thickness. The most disagreement was observed in soil frost classes 6-10 and 11-20 when there is no snow cover present in the ground. Here the percentages of disagreeing predictions were over 50 %. Similar over 50 percent inconsistency was observable in snow classes 6-10 and 11-20 during thaw ground. On average 33 % of these predictions disagreed during frozen ground. Overall the FT-ESDR and ERA-Interim agreed better during thick soil frost and snow cover.

Table 18. Error (%) rate of observations per frost/snow class. Colour gradient: orange (high) – white (low). FT-ESDR=Freeze/ Thaw Earth System Data Record. SYKE=Suomen Ympäristökeskus (Finnish Environmental Institute). ST1=Soil Temp level 1 (7 cm).

SYKE IN-SITU	FT-ESDR SNOW (cm)						ERA Interim ST1 SNOW (cm)					
SOIL FROST (cm)	0	1-5	6-10	11-20	>20	Mean	0	1-5	6-10	11-20	>20	Mean
0	17	32	57	46	83	47	15	22	22	41	75	35
1-5	76	51	39	26	11	41	68	47	37	26	12	38
6-10	65	39	34	30	17	37	63	28	25	15	8	28
11-20	71	44	38	14	13	36	43	23	19	12	8	21
>20	60	29	25	16	5	27	50	8	16	8	4	17
Mean	68	41	34	22	12	35	56	27	24	15	8	26

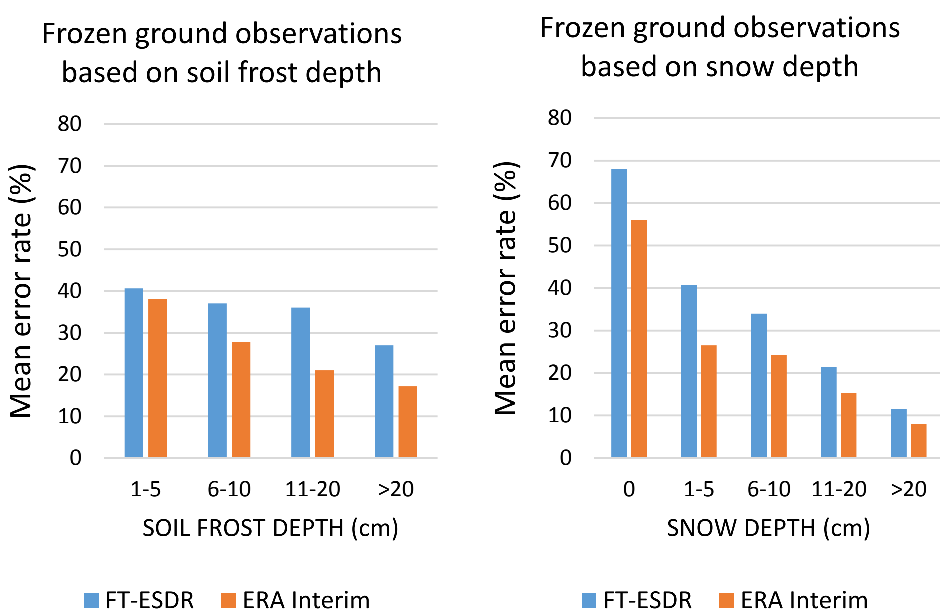


Figure 28. Error rates of FT-ESDR predictions and ERA Interim reanalyse with SYKE in-situ observations. FT-ESDR=Freeze/ Thaw Earth System Data Record. SYKE=Suomen Ympäristökeskus (Finnish Environmental Institute).

Table 19. Percentage (%) of disagreeing predictions between FT-ESDR and ERA Interim. Colour gradient: orange (high) – white (low). FT-ESDR=Freeze/ Thaw Earth System Data Record. SYKE=Suomen Ympäristökeskus (Finnish Environmental Institute). ST1=Soil Temp level 1 (7 cm).

		ERA Interim ST1					
		SNOW (cm)					
		0	1-5	6-10	11-20	>20	Mean
FT-ESDR	0	20	38	51	50	34	39
SOIL FROST (cm)	1-5	46	46	46	37	17	38
	6-10	53	42	37	31	17	36
	11-20	55	40	39	20	18	34
	>20	30	34	29	20	7	24
	Mean	46	41	38	27	15	33

7. Discussion

7.1 Multivariate analysis of soil frost

In the GAM output of open areas, the air temperature and snow cover interaction plot showed that the increasing snow layer diminished the effect of air temperature. When this snow layer reached a thickness of 40-45 cm its effect to the response becomes less important. This is corresponding with the literature where Williams & Smith (1989) stated that snow layer become less effective after 50-60 cm thickness. In the LME model, the air temperature had a more severe decreasing effect to frost layer below (−4.5) than over (−3.3) 45 centimetres snow layer on open areas. This is indicating that increasing snow layer is diminishing the effect of air temperature as expected. The snow layers had decreasing effect to soil frost (−0.7) during ≤45 cm snow cover. This effect were clearly smaller (−0.2) when the snow layer was >45 cm.

In the GAM results of forest model, the interaction and threshold were different on forests because of tree covers. As expected based on literature and previous studies, the GAM interaction plot showed that snow layers effect on response is less significant in forested areas than on open areas which is most likely due to canopy (Mustonen 1986; Venäläinen et al.

2001b). However, there were some indications that for a thin snow layer (<20 cm) and the air temperature below −2 °C, this would have a positive effect on our response. The possibility of this increasing effect of thin snow cover at the beginning of freezing season was also pointed out in literature by Williams and Smith (1989). From LME model we can notice that the effect of snow cover to response is −0.1 per one unit.

For bogs GAM interaction plot, the snow layer had an increasing effect toward our response until snow layer depth reached 30 cm thickness. The LME model values for bog sites showed that the effect of one degree change in air temperature to soil frost depth (−1.6 cm) was the smallest from all three sites which was expected since the peat material in soils has a good insulating effect and the soil in bogs has typically higher moisture content than soil in open and forested areas (Williams & Smith 1989). Surprisingly LME model is indicating that a snow layer has a positive effect (0.4) to soil frost thickness if the snow layer depth is less than 30 cm. For thickness of over 30 cm, the effect turns to negative (−0.2). In literature, this early thin snow layer has been linked with the positive cooling effect to ground and soil frost formation until it reaches sufficient insulating thickness (Yershov 2004: 361-366, Williams & Smith 1989).

Based on chosen explanatory variables in GAM analysis, monthly mean precipitations (*AVE.Prec*) showed clearly a negative effect towards our response between 0 and 45 mm and after this, it stabilized. However variable p-value suggests that the effect is not as strong in forest and bog models. LME gave a negative value to seasons monthly mean precipitation (*AVE.Prec*) in the situation below 45 mm threshold on open (−0.2 cm), forest (−0.5 cm) and bog (−0.2 cm). The situation where there have been less than 45 mm rainfall showed to lead to a clear negative effect on all three site types. The estimated effects of precipitation, corresponded well with GAM smooth plots where there was a noticeable negative effect with the small values of precipitation in all three models (Figure 3, 8 & 13). The interaction terms showed that when we pass this limit the effect turn around to positive on open (0.1) and forest (0.09). On bog, the effect was still negative (−0.03) however in all cases the effect were small staying close to zero. Over 45 mm rainfall showed to lead to a clear negative effect on all three site types.

Surprisingly the freezing index was not statistically significant in any of our GAM or LME models. In previous studies considering frozen ground, the freezing index has been one of the most significant variables when determining soil frost depth (Seppälä 1999; Frauenfeld et al. 2004; Frauenfeld & Zhang 2011; Shiklomanov 2012). However, in several cases the research has been carried out on snow free areas (Venäläinen et al. 2001b; Gregow et al. 2011b) excluding snow cover and its effect. However, William & Smith (1989) pointed out that soil frost depth cannot sufficiently correlate with wintertime cooling. In our study, freezing index did not show sufficient statistical evidence (nor $p < 0.05$ neither $p < 0.1$) in any of our three GAM or LME models.

The NAO index was statistically significant in all three GAM models and on open area and forest LME models. The GAM smooth plots indicated that it does not have clear increasing nor decreasing effect but overall soil frost maximum depth should be higher when the NAO index value is close to zero (Figures 5, 10 & 15). In the LME models, the NAO index quadratic terms implied that highest depths in soil frost layer is reached when peak around zero for open (0.05 index value) and forest (0.11 index value). These were corresponding with GAM smooth functions. The NAO index did not show sufficient statistical evidence in bog model.

The spatial variables (*coord.east*, *coord.north*) and more specifically the north coordinate had an increasing effect to soil frost depth (Figures 6 & 11). There was overall less distinct differences between frost tubes in northern Finland on bog environment compared to open and forest (Appendix 10). The bog models coordinate plot (Figure 16) had clearly smaller interval between interaction plot contour line values compared to open and forest (Figure 7 & 11). This might be the reason why soil frost maximum thickness showed to increase southwards which was complete opposite compared to open and forest. These observations are implying that spatial variables, especially north-south, does not affect soil frost as much on bog sites. In LME models the results with spatial variables also showed that the evolvement of soil frost is highly dependent on the location. When these spatial variables were considered separately on LME models, the north coordinate were statistically significant and implied that the maximum depth of soil frost would increase by 2 cm on open, 3 cm on forest and decrease by 1 cm on bog when moving 100 km northwards. The east coordinate did not show sufficient statistical evidence in any of LME models.

As expected based on previous studies, soil frost maximum decreases eastwards and increases northwards (Soveri & Varjo 1977; Huttunen & Soveri 1993: 71; Solantie 1998; Orvomaa 2015). SYKE's report 2015 about frozen ground stated that Finland should be divided into two main soil frost areas instead of the previous five areas suggested by Soveri and Varjo (1977) in their research about soil frost depths in Finland. The presence of the sea and its warming effect diminish average soil frost depths nearby coast compared to inland sites and lead to considerably different conditions compared to northern Finland. This makes it more reasonable to consider south and north as own study areas. Northern Finland average soil frost layer thickness is clearly higher on northern measurement sites compared to southern sites. However, based on GAM and LME models the air temperature and north coordinate were both statistically significant. Hence this is indicating that some other factor, which we did not take into account, is also affecting soil frost depth and leading to these differences between northern and southern Finland.

The year variable (*season.id*) showed a decreasing trend during our measurement period in GAM models even though it was deemed as statistically significant only in forest GAM and LME models. In LME models the same forest variable indicated a declining trend where maximum depth of soil frost would decrease by 0.17 cm per year. This result supports GAM analysis results and previous studies about the projected diminishing extent of the seasonally frozen ground area (Venäläinen et al. 2001b; Henry 2008; Frauenfeld & Zhang 2011).

The *site* variable which was modelled as a random effect to our GAM models was highly important based on random effect outputs on all three models. Due to this we were able to take into account the differences between frost tube locations in Finland. In LME models the same *site* variables were random effect in each cases. The results showed that the differences in the location of frost tube are significant factors when we are determining soil frost layer thicknesses on the regional scale. The standard deviation on open (15.0) and forest (16.6) were almost equal and on bog (7.8) this value was clearly smaller.

The results of GAM models corresponded well with our presumptions and with the previous research (Liu et al. 2010) by showing the importance of air temperature and snow layer as the main factors in each environment types. However we assumed that freezing index would have been more important explanatory factor when determining soil frost layer maximum thickness as suggested in previous studies (Soveri & Varjo 1975; William & Smith 1989; Seppälä 1999; Salonen 2002; Frauenfeld et al. 2004; Frauenfeld et al. 2007, Shiklomanov 2012). In our study the freezing index variable did not show any statistically clear effect towards response. Several previous studies have shown that there is a strong correlation between the annual freezing index and the maximum of the soil frost depth (Tomoyoshi et al. 2006) when snow cover is reasonably shallow. However, like Tomoyoshi et al. (2006) stated in their study, after the snow layer reaches a certain threshold the insulating effect is diminishing the effect of freezing index to the soil. This may be a plausible reason why we are not observing any statistically significant correlation between *FI* and response. Nevertheless, the interaction plots were consistent with previous studies and showed that the insulating effect of the snow covers is not linear (Henry 2008: 430). Also, the difference between the three types of frost tube sites was more distinct than anticipated. NAO –index had a negative effect on a response during both positive and negative phase on open areas as Frauenfeld & Zhang (2011) studies implied. However, this was evident only on forest and bog sites.

The interaction plots of GAM confirmed our presumptions that the effect of air temperature to the soil frost depth is dependent on the snow layer thickness and the interaction varies depending on the environment (open, forest and bog). Interaction plots from open, forest and bog sites indicated that snow cover will affect to response until it reaches a certain threshold, after which its effect diminishes. LME analysis also supported

GAM smooth plots and our presumptions considering snow and air temperature interaction and their effect to frozen ground. Increasing snow layer diminishes air temperature effect and in certain point insulates the ground. The results also showed that there are noticeable differences between three environmentally different frost observation sites (open, forest & bog). Canopy and its shading effect to the ground lead to shallower snow cover and different conditions for soil frost formation compared to open areas. However, there were less distinct differences between open and forest sites compared to bog. Most likely organic peat soil lead to significantly different circumstances in ground properties in similar climate. On open and forest sites there have been usually installed two frost tubes and on bog just one. Due to this, the average of bog sites is done with the lower number of observations which overall leads to higher error variance.

From residual plots of GAM models (Appendix 6a-c), we can notice that residuals are spreading out when moving towards to right on the x-axis. This is indicating that error variance is increasing when the fit is increasing and is causing so-called heteroscedasticity of error terms on our data. However, this should not affect regression coefficients estimates (Molenberghs & Verbeke 2001). The repeated yearly basis, measurements from the same frost tube sites are also causing autocorrelation error in analysis. This could've resolved by taking the 30 year average of maximum soil frost layer thickness for each site. However, in our tests, this did not give any significant changes and continuing with yearly data was a more reasonable approach. The secular autocorrelation was reduced by changing monthly data to annual average which reduces the effect of autoregression. This means that we cannot make any reliable interpretation based on variables *gam()* function p-value. Another thing is that because data contains also zero values, the model fits some of these as a negative value as we can see from on fitted values plot (Appendix 6a, 6b & 6c). This is more noticeable and bigger problem on the last bog model where there is clearly more zero values. These factors may affect the reliability of the obtained p-values.

We tried to fix these problems by using logarithmic transformation to our response and also transform zeroes to small positive values and see if this affect to residuals distribution and fitted values. However, this approach did not solve the problem. With the yearly level data, we cannot take into account all possible monthly changes and therefore this is likely generalising too much analysis. However, this autocorrelation is not affecting considerably the consistency of model estimates (Molenberghs & Verbeke 2001). To be able to take into account these problems we would need to change the used statistical model.

SYKE's report about soil frost depths 1981-2010 and the differences between northern and southern Finland (Orvomaa 2015) suggest that we should analysis south and north frost tubes separately. Due to GAM results and significance of site variable, it is justifiable to state that there are distinct differences of frozen ground formation on different parts of Finland. We tested as a separate analysis for north and south frost tubes in same GAM analysis. This test did not show significantly different results in interaction plots on open and forest sites. However, on bog this led to very different results which could be due to the high amount of zero values in the southern part of data. Overall the average soil frost depth in northern Finland is higher compared to southern Finland.

In this study, several variables were extracted from rasters with 10x10 km spatial resolution. This reasonably large cell size has most likely generalized too much the actual conditions and affected to results. Huttunen and Soveri (1993) pointed out, that more detailed information on auxiliary factors such as soil moisture and groundwater level are required to make more accurate estimates on the evolvement of the soil frost layer on the regional scale. In addition, ground temperature would tell us more about the proceeding of the freezing process during the freezing season and therefore lead to more reliable results. Properties like grain size and the portion of mineral land organic material are affecting soil thermal and physical properties and in the end timing and duration of frozen ground. Soil properties clearly have an effect to soil frost maximum depth which would suggest need to use soil samples from the soil frost observation sites.

Frozen ground depth can vary significantly in short distances due to unnoticeable differences on the soil surface and forecasting this based on climate is inaccurate (Williams & Smith 1989: 15-17). For example, we have no information about the level of possible groundwater on the frost observation sites which can be diminishing soil frost depth in some of these sites. Due to differences on soil properties, we cannot precisely forecast soil frost distribution and depth and predictions about this can mostly be estimated in heterogeneous soil types (Williams & Smith 1989; Liu et al. 2010). In the further inspection of data, there were a relatively high amount of cases where frost layer bottom limit was measured but the top limit was not. The number of observations starts around 1999 and count of these rises on average 580 measurements per site 1981-2010. However, in these cases where the bottom layer is deeper than the couple centimetres during autumn we could make an assumption that the top layer depth is around zero. We tested situation where we included these observations by setting all these cases top depth as 0 cm during autumn where the bottom depth had been measured but the top was marked as NULL. The test did not change the snow and air temperature interaction estimated smooth function significantly in any of three sites but the effect of other variables diminished. Nevertheless, because we have no certain information about the level of frost layer top we are forced to exclude these observations from the analysis. This creates significant gaps in our frost season study sample. In further analysis it would be recommended to include information about soil moisture, temperature, and groundwater level to get a comprehensive view of the frozen ground. The GAM analysis results were consistent with our presumptions and previous research about variables which are determining soil frost layer thickness. The LME analysis supported GAM results. Overall these results gave us more insight about the total effect of the explanatory variables to frozen ground.

7.2 Northern Finland soil frost tubes time series trend analysis

In the first part of our results in Figures 18-20, there were some noticeable differences in the average of maximum soil frost layer thickness, not just between locations but between environmental types, as well. On open area sites like Inari Angel and Sodankylä Tähtelä soil frost averages were clearly higher than environment overall mean. On the other hand average of Kemijärvi and Kittilä were lower than it. The site specific variance of yearly maximum of soil frost varied less on forests than on open areas. On forest sites, in addition to Inari Angel and Sodankylä Tähtelä, also Inari Nellim and Utsjoki averages were over the mean. Means in Kemijärvi, Enontekiö Iitto and Enonteiö Kilpisjärvi were clearly below the overall mean on forest sites. From all three environment types, the bog sites differed from the others with distinct lower averages on all sites which were just one fourth (15 cm) of the two others (60 cm). The greatest soil frost depths on bog sites were on Inari A and lowest on Enontekiö Iitto, Kilpisjärvi and Kemijärvi.

The soil type of each frost tube might explain some of these differences between the distinct locations. In Appendixes 1, 2 and 11 we have illustrated the soil type descriptions and mean of maximum soil frost layer thickness in different soil types on open and forest sites. On sites where the soil type is fine-grained clay (Sa) (Enontekiö Kilpisjärvi., Kittilä) or coarse-grained moraine (Mr) (Kemijärvi), the soil frost depths tend to be clearly shallower compared to others sites (see Figures 18 & 19). These results are supporting previous literature and studies (Seppälä 1999; Yershow 2004) view that ground freezes deeper on coarse material like sand and is shallower on fine material like clay. This is linked to soil type and its material thermal properties as heat capacity and thermal conductivity of the ground (Mustonen 1986: 94-98: Williams & Smith 1989: 50-100; 94-98; Seppälä 1999; Salonen et al. 2002: 100-119). On bogs, the ground organic material varies between different kinds of peat types which we have excluded from this analysis. Based on the Figures 18-20 generally, the soil frost layer appears to be thinner on peat soil. Mustonen (1986) stated that organic materials on average froze less mainly because of differences in material heat capacity, thermal conductivity, and different conditions on the ground. The greater mass heat capacity in peat soil results to greater amount of stored heat during summer which slows ground cooling in autumn. Even though peat ground froze faster pace than mineral soil (Yershow 2004) and the thermal conductivity of saturated and frozen peat is relative high, overall soil frost reaches lower depth on peat material based on this study results. In contrary for example at Inari Angel open site, ground material is fine sand and on forest finer sand. Both of these can be considered as a frost susceptible coarse materials (Yershov 2004: 358). However if soil material becomes too coarse, average frost layer diminishes like for example in cases where the soil contains moraine (Mr). On the other hand in Kittilä's open site soil type is clay which tends to have lower soil frost layer compared to others due to its fine texture. As previous studies (Soveri & Varjo 1977; Mustonen 1986; Huttunen & Soveri 1993) have stated the greatest frost depths are reached on coarse materials. In our study, these were sand (HHk), gravelly sand (SrHk) and silty sand (HtHk) which grain size varies 0.02-20 millimetres.

When we considered maximum depth changes of soil frost between different months we noticed that there have been some major declines especially on open and forest. These declines are clearly focused on the middle and end of frost season. On open sites there have been statistically significant (α=95 %) decrease on soil maximum depth between January-May and on forest sites January-June in northern Finland. This decrease has been on these months on open areas on average 2 cm per year and on the forest on average 2.46 cm per year. In addition to these, there has been a decrease in December both open (1.48) and forest (1.82) site. Again bog as an environment has experienced distinctly fewer changes and there was no significant (α=95 %) decrease in any month. However, the bog sites average soil frost depths were significantly lower compared to open and forest so overall decrease could be expected to be lower as well. The distinct differences in soil material and conditions in bog sites are most likely leading so clear differences on soil frost depth. Results are overall supporting and consistent with the previous estimates about decreasing snow cover and frost season during springtime (Venäläinen et al. 2001b; Déry & Brown 2007; Sutinen et al. 2008).

Venäläinen et al. (2001a) and Jylhä et al. (2009) predicted that soil frost will be shallower and frost season will be shorter in Finland in the future. Our study results support these and other studies (Tomoyoshi et al. 2006; Jylhä et al. 2008; Sutinen et al. 2008; Henry 2008; Frauenfeld & Zhang 2011; Bilotta et al. 2015) about shortening frost season. We can see from Figure 24 and Table 15 that there has been a clear decline in frost season length. Climate change has led to an increase in mean air temperature and annual precipitation especially during winter in high latitudes. Due to this wintertime freezing index has decreased and in the other hand thawing index increased (Williams & Smith 1989, Peng et al. 2013). This has led to shorter frost season and overall shallower soil frost depths in the past 50 years (Frauenfeld & Zhang 2011; Gregow et al. 2011a). The extending thawing depth during spring and changes in snowpack have been also linked to reduced carbon emitting capability of northern hardwood forest (Groffman et al. 2006) and increasing greenhouse emissions from the ground (Heyer et al. 2002; Öquist & Laudon 2008). There have been estimates which indicates that this kind of diminishing of seasonal soil frost depth and overall frost season length will continue in Finland in the future (Venäläinen et al. 2001b; Gregow 2013). Frozen ground will be rarer and less deep in southern Finland and possible increasing snow cover will diminish soil frost depth in northern stations (Venäläinen et al. 2001b). The soil frost maximum thickness and snow cover decline was observed on all three site types for a period of 1981-2010 in northern Finland. Based on Figures 21-23 the decline seems to be greatest on forest and clearly lower on bog sites. Statistical trend tests supported this interpretation based on Table 16. However, Yershov (2004: 348) pointed out that such a long-term variation on soil frost maximum freezing and thawing depth is not uncommon.

In some locations, the changes have been more severe over past decades based on our results. For instance, all bog sites showed just a minor decline compared to open and forest sites where the decline of maximum frost depth was at least two times higher. Overall largest decline was experienced in northern Finland sites where average of soil frost maximum depth was clearly over the mean. The greater soil frost depth in the forest could be explained with shallower snow cover and weaker insulating effect due to overhead canopy as Mustonen (1986) and Venäläinen et al. (2001b) pointed out. The tree cover is also affecting forest ground floor overall same way as an insulator as snow cover, is accumulating snow and keeping snow layer beneath trees shallower. On the other hand, there is no significant difference on the maximum snow depth on open areas and forests.

Nevertheless, changes on the frost season length were almost equal on three site types so the overall duration of frozen ground was not dependent on the environment. These results are consistent with previous researches and are supporting the previous assessments and predictions which due to increasing winter air temperature, have predicted declining trend on frozen layer thickness and frost season length in Finland and globally (Venäläinen et al. 2001b; Tomoyoshi et al. 2006; Jylhä et al. 2008; Sutinen et al. 2008; Gregow et al. 2011b; Frauenfeld & Zhang 2011; Luo et al. 2017).

The original SYKE frost tube data was sorted from daily measurements to monthly values which were used to calculate yearly maximums for each site in northern Finland. The possible source of errors in the analysis are most likely due to this sorting of original data or excluding observations as was mentioned in GAM analysis where frost layer top measurement was missing. The exact beginning of the frost season especially during autumn can vary when due to air temperature fluctuation around 0 °C degree.

7.3 Correspondence of in-situ observations against satellite measurements

The comparison of NASA's FT-ESDR landscape state predictions and ECMWF's ERA-Interim soil temperature reanalyse values with SYKE's soil frost in-situ measurements in Finland showed some clear and also expected differences on estimation accuracy for different soil frost and snow depth conditions. Kim et al. (2011) stated in their research that FT-ESDR has its limitations especially considering conditions under thick snowpack and ground vegetation.

As expected the estimation accuracy steadily increases with thickening snow cover and soil frost layer (Table 18). However, the error rate stayed clearly over 50 % during zero snow cover even if frozen ground layer reached over 20 cm thickness. These results are indicating that the evaluation process behind FT-ESDR product might actually use snow cover as an indicator for predicting soil frost occurrence. It is also possible that the remote sensing based estimates are more sensitive to increasing snow layer than actual freezing of the ground. Another explanation could be the result of increased soil moisture content in the ground during autumn. At the beginning of freezing season the first snow layers are likely thaw before the ground gets more permanent snow cover. Here the moisture content of melting snow layer is increasing which diminishes microwave sensor ability to reach the ground and evaluate the state of it even though the frozen layer might be over 10 cm thick. Kim et al. (2017) pointed out that the 37 GHz microwave sensor is optimal under dry conditions and increasing moisture on the soil surface and ground vegetation are likely limiting its ability to make reliable predictions about the actual state of the ground. ERA-Interim seemed to match better with in-situ observations. In Table 18 and Figure 30 are presented the error rates for each soil frost/snow class. The error rates are highest on snow class 0 for both FT-ESDR and ERA-Interim. However, the thaw ground observations error's percent implies that estimation accuracy is diminishing with increasing snow layer. This effect of vegetation and snow cover is leading shallower microwave emitting depth as was mentioned in Kim et al. research (2017). Overall ERA-Interim is matching better with frozen ground in-situ observations compared to FT-ESDR (Figure 30). The mean error rate of estimates were on FT-ESDR (PM/AM) around 29 % and ERA-Interim (ST1) around 20 % (Appendix 4 and 5). Nevertheless, the overall behaviour is similar in different frost and snow classes. The accuracy of observations are getting better with increasing snow layer and soil frost. Overall results are implying that there is a clear difference between FT-ESDR and ERA-Interim in certain snow and soil frost conditions. However, during a thicker soil frost layer or shallow snow layer, the values were clearly more consistent with in-situ observations and with each other. However, it should be keep in mind, that even though these products can be both used in frozen ground studies, they use different methods to create their estimates. FT-ESDR is based on microwave measurements in its landscape freeze/thaw state estimates whereas ERA Interim uses air temperature and atmospheric model to estimate soil temperature.

These results about FT-ESDR accuracy are consistent with previous studies () which emphasized the difficulty on estimation during seasonal transition periods. Our study consists three observation per month and this annual time period consist of months from September to December. This period includes observations from the end of thawing period at the beginning of freezing season. The FT-ESDR overall mean of accuracy during our study period was approximately 72 % and on ERA-Interim around 80 %. FT-ESDR evaluations (Kim et al. 2017) were calibrated with surface air temperature values from several stations around the world which gave accuracy rate within 96-80 % (PM overpass) and 89-73 % (AM overpass) in Northern Hemisphere and these rates tended to decrease from September to December. Compared to these, our results with in-situ observations led to lower accuracy rates with FT-ESDR predictions during autumn in Finland. ERA-Interim error rates were slightly lower than FT-ESDR and seemed to match better with in-situ observations especially during shallow snow and soil frost. It was pointed out by Albergel et al. (2015) that there is uncertainties with ERA Interim ability to estimate soil temperature under snow cover. Also the representativeness of the estimate for entire 16x16 km cell is problematic especially during snow cover.

Hence the results are showing that the accuracies of the FT-ESDR estimate and ECMWF ERA-Interim model are in overall reasonable considering the difficult conditions which often varies significantly especially during seasonal transitional periods. The spatial resolution of FT-ESDR product and ERA-Interim model are rather coarse. Comparison to point scale soil frost observations is challenging. The ways to increase the accuracy during transition seasons like autumn and spring would require more information about ground properties like moisture content and temperature.

8. Summary and conclusions

In this study, we had three major objectives. 1) Get a better and more comprehensive picture of the formation of frozen ground in different kinds of conditions and inspect the key explanatory variables behind this phenomenon. 2) Constitute an outlook about soil frost depth and frost season length changes and current trends for the period from 1981 to 2010 in Finland. 3) Examine the ability of in-situ observations to validate the accuracy of the nowadays common remote sensing product (FT-ESDR) within regional scale in Finland. For reference also ECMWF ERA-Interim analysis model estimates for soil temperature was evaluated. The research of all these three sections is based on soil frost observation records 1981-2010 in Finland conducted by SYKE. Remote sensing evaluations about soil freeze/thaw state and temperature data were provided by NASA and ECMWF databases. On data collection and for variable arrangement we used geographic information systems (GIS), statistical software's and open access remote sensing data. The analysis was carried out with R statistical software version 3.5.0 utilizing statistical models like generalized additive-, linear mixed-effect models and contingency tables from tool packages *mgcv*, *nlme* and *gmodels*. In addition, Salmi et al. (2002) MAKSENS Excel template and its Mann-Kendall and Sen's slope estimate tests were used.

Multivariate analysis was performed with generalized additive model and extended with the linear mixed-effects model. The results revealed that variables such as precipitation, air temperature, snow maximum depth and spatial location of frost tube are significantly affecting to frost season and the maximum thickness of soil frost layer in Finland. Interaction plots showed that the effect of air temperature was diminishing with increasing snow depth. For example, after the snow layer reaches about certain threshold on open areas, the insulating effect become stable and any increase on snow layer are clearly less important. However, results indicated that the differences in soil frost depths between northern and southern Finland cannot only be explained with air temperature. Statistical significance of north coordinate and air temperature variables was Plots of the estimated smooth function revealed that this interaction differs on open areas, forest and bog sites which is likely due to soil type and canopy.

Overall there were distinct differences on soil frost depth between northern and southern Finland. Trend analysis about northern Finland frost tube observations indicated that soil frost maximum depth has overall decreased in period 1981-2010. This decrease has been clearly more severe in certain environments and frost tube sites in northern Finland. These differences can be mostly explained with differences in soil type thermal and physical properties. The overall length of frost season has decreased similarly on open, forest and bog sites during our study period.

The FT-ESDR landscape freeze/thaw state and ERA-Interim soil temperature product estimate consistency with in-situ observations in Finland varied. These were highly dependent on current snow cover and frozen ground thickness. FT-ESDR overall mean of error rate (35 %) was nine percent points higher compared to ERA-Interim (26 %) but overall the behaviour of accuracies was similar in different classes. As expected, the satellites ability to observe thaw ground decreased with increasing snow cover. This is most likely linked with the moisture content of snow and soil surface during autumn months. The results of this study are supporting previous research on the subject. The air temperature and snow cover are ones of the most significant variables when we are considering the overall soil frost depth in the different types of environments.

Due to characteristics of the frozen ground system which are highly dependent on local factors, it is difficult to create accurate predictions about soil frost depth in regional scale. The effect of different explanatory factors to the frozen ground is not always linear and these tend to have mutual interactions which in the other hand have an effect on the frozen ground. In order to create more precise results, additional in-situ measurements from frost tube sites such as soil moisture and temperature would be needed. The distinct differences between open, forest and bog sites and the significance of soil type to the annual maximum depth of soil frost are suggesting that we should consider conducting the analysis to more homogenous environments. Because of the seasonality of the phenomenon, the use of monthly or even daily data should be utilized. However, the current sources of errors such as the autocorrelation and heteroscedasticity of residuals should be taken into account more carefully in the future. The possible way to overcome these problems would most likely require modifying of the model so it would take into account autovariance and use for example different distribution than Gaussian.

Our study aimed to give more insight about frozen ground and these multidimensional processes behind it and past changes in Finland. The current results could be applied in the planning of future researches when considering studies configuration. Overall we can say that the multivariate models as GAM and LME can be useful tools when we are investigating complex environmental phenomenon as frozen ground. Especially in cases when multiple different variables need to be taken into account. With long-term observations, it is possible to identify recent trends of frozen ground on the regional scale. As Brown et al. (2000) stated, yearly and monthly observations about soil frost depth in regional and hemisphere-scale are needed to be able to understand possible impacts of climate changes to cold ground and permafrost in the future. The vaster material from a broader area and a longer period of time could help us to understand phenomenon also on a global scale. The remote sensing methods have proved to be useful tools for evaluating frozen ground on a large scale (Han et al. 2014). The continuing development of these and increasing observations from larger scale are going to help upcoming researches. The better understanding of seasonally frozen ground is going to have a significant impact considering environmental, economic and social aspects in the future.

9. Acknowledgements

I would like to thank all my supervisors for all the help and guidance, which they have provided along the whole study process. Thanks are due to professor Jan Hjort from Uni. Oulu Geographic Research Unit for all the counselling and advice and to doctor Kimmo Rautiainen from Finnish Meteorological Institute for all the help with the topic and data processing. Also many thanks to doctor Leena Ruha from Uni. Oulu Research Unit of Mathematics Sciences for all the valuable help with statistical analysis. I'll want to thank the Finnish Meteorological Institute and Arctic Space Centre for the possibility to write my thesis and for providing financial backing. In addition, thanks to every FMI personnel who have helped me during this process especially the people in Tähtelä. Lastly, I want to thank my family and friends for all the support during my years at the university.

10. References

Aaltonen, V. T., Aarnio, B., Hyyppä, E., Kaitera, P., Keso, L., Kivinen, E., Kokkonen, P., Kotilainen, M., Sauramo, M., Tuovila, P. & J. Vuorinen (1949). Maaperäsanaston ja maalajien luokituksen tarkastus v. 1949. Summary: A critical review of soil terminology and soil classification in Finland in the year 1949. Maataloustieteellinen aikakauskirja 21, 37-66.

Albergel, C., Dutra, E., Muñoz-Sabater, J., Haiden, T., Balsamo, G., Beljaars, A., L., Isaksen, P., de Rosnay, I., Sandu & N. Wedi (2015). Soil temperature at ECMWF: An assessment using ground-based observations. *Journal of Geophysical Research*: *Atmospheres*, 120(4), 1361-1373.

Allison, I., R. G. Barry & B. E. Goodison (2001). *Climate and Cryosphere (CliC) Project Science and Co-ordination Plan: Version 1* (Vol. 114). Joint Planning Staff for WCRP, World Meteorological Organization. 76 p.

Bayard, D., M. Stähli, A. Parriaux & H. Flühler (2005). The influence of seasonally frozen soil on the snowmelt runoff at two alpine sites in southern Switzerland. *Journal of Hydrology, 309*(1), 66-84.

Bilotta, R., Bell, J. E., Shepherd, E., & A. Arguez (2015). Calculation and evaluation of an air-freezing index for the 1981–2010 climate normals period in the coterminous United States. *Journal of Applied Meteorology and Climatology, 54*(1), 69-76.

Bojariu, R., R. García-Herrera, L. Gimeno, T. Zhang & O. W. Frauenfeld (2008). Cryosphere – Atmosphere interaction related to variability and change of northern hemisphere annular mode. *Annals of the New York Academy of Sciences, 1146*(1), 50-59.

Bondell, H. D., A. Krishna & S. K. Ghosh (2010). Joint variable selection for fixed and random effects in linear mixed-effects models. *Biometrics*, *66*(4), 1069-1077.

Brown, J., Hinkel, K. M., & F.E. Nelson (2000). The circumpolar active layer monitoring (CALM) program: research designs and initial results. *Polar geography*, *24*(3), 166-258.

Campbell, J. L., S. V. Ollinger, G. N. Flerchinger, H. Wicklein, K. Hayhoe & A. S Bailey (2010). Past and projected future changes in snowpack and soil frost at the hubbard brook experimental forest, new hampshire, USA. *Hydrological Processes, 24*(17), 2465-2480.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. & F. Vitart (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553-597.

DeGaetano, A. T., M. D. Cameron & D. S. Wilks (2001). Physical Simulation of Maximum Seasonal Soil Freezing Depth in the United States using Routine Weather Observations." *Journal of Applied Meteorology* 40.3 (2001): 546 p.

Déry, S. J., & R. D. Brown (2007). Recent Northern Hemisphere snow cover extent trends and implications for the snow-albedo feedback. *Geophysical Research Letters*, *34*(22), 1-6.

Diaz, H. F., Eischeid, J. K., Duncan, C., & R.S Bradley (2003). Variability of freezing levels, melting season indicators, and snow cover for selected high-elevation and continental regions in the last 50 years. In *Climate Variability and Change in High Elevation Regions: Past, Present & Future.* Springer, Dordrecht. 33-52.

Drebs, A., Nordlund, A., Karlsson, P., Helminen, J. & P. Rissanen (2002). Tilastoja Suomen ilmastosta 1971–2000. Ilmastotilastoja Suomesta, 2002:1, 99 p.

Fienberg, S. E. (2007). *The analysis of cross-classified categorical data*. 198. Springer Science & Business Media.

Frauenfeld, O. W., Zhang, T., Barry, R. G., & D. Gilichinsky (2004). Interdecadal changes in seasonal freeze and thaw depths in Russia. *Journal of Geophysical Research: Atmospheres, 109*(D5).

Frauenfeld, O. W., T. Zhang & J. L. Mccreight (2007). Northern hemisphere freezing/thawing index variations over the twentieth century. *International Journal of Climatology, 27*(1), 47-63.

Frauenfeld, O. W., & T. Zhang (2011). An observational 71-year history of seasonally frozen ground changes in the Eurasian high latitudes. *Environmental Research Letters*, *6*(4).

French, H. M. (2017). *The periglacial environment* (4th ed.). 515 p. Wiley-Blackwell.

Gilbert, R. O. (1987). *Statistical methods for environmental pollution monitoring*. 321 p. New York: Van Nostrand Reinhold.

Gorham, E. (1991). Northern peatlands: role in the carbon cycle and probable responses to climatic warming. *Ecological applications*, *1*(2), 182-195.

Gregow, H., H. Peltola, M. Laapas, S. Saku & A. Venäläinen (2011a). Combined occurrence of wind, snow loading and soil frost with implications for risks to forestry in Finland under the current and changing climatic conditions. *Silva Fennica, 45*(1), 35-54.

Gregow, H., K. Ruosteenoja, I. Juga, S. Näsman, M. Mäkelä, M. Laapas & K. Jylhä (2011b). *Lumettoman maan routaolojen mallintaminen ja ennustettavuus muuttuvassa ilmastossa*, Ilmatieteen laitos. Raportti 2011:5. 45 p.

Gregow, H. (2013). *Impact of strong winds, heavy snow loads and soil frost conditions on the risks to forests in northern europe.* Ilmatieteen laitos. 44 p.

Groffman, P. M., J. P. Hardy, C. T. Driscoll & T. J. Fahey (2006). Snow depth, soil freezing, and fluxes of carbon dioxide, nitrous oxide and methane in a northern hardwood forest. *Global Change Biology*, *12*(9), 1748-1760.

Guisan, A., T. C. Edwards & T. Hastie (2002). Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling, 157*(2), 89-100.

Han, L., A. Tsunekawa, M. Tsubo, C. He & M. Shen (2014). Spatial variations in snow cover and seasonally frozen ground over northern china and Mongolia, 1988–2010. *Global and Planetary Change, 116*, 139-148.

Hardy, J. P., P. M. Groffman, R. D. Fitzhugh, K. S. Henry, A.T. Welman, J. D. Demers, T. J. Fahey, C. T. Driscoll, G. T. Tierney & S. Nolan (2001). Snow depth manipulation and its influence on soil frost and water dynamics in a northern hardwood forest. *Biogeochemistry*, *56*(2), 151-174.

Hastie, T. & R. Tibshirani (1990). *Generalized additive models*. 335 p. Chapman & Hall.

Heino, R., & E. Hellsten (1983). *Tilastoja Suomen ilmastosta 1961-1980*. Ilmatieteen laitos.

Hentilä, V., E. Siunga & A. Palolahti (1994). *Roudan vaikutusten mallintaminen.* Tielaitoksen selvityksiä. 81 p.

Henry, H. A. (2007). Soil freeze–thaw cycle experiments: trends, methodological weaknesses and suggested improvements. *Soil Biology and Biochemistry*, *39*(5), 977-986.

Henry, H. A. (2008). Climate change and soil freezing dynamics: historical trends and projected changes. *Climatic Change*, *87*(3-4), 421-434.

Heyer, J., U. Berger, I. L. Kuzin & O. N. Yakovlev (2002). Methane emissions from different ecosystem structures of the subarctic tundra in Western Siberia during midsummer and during the thawing period. *Tellus B: Chemical and Physical Meteorology*, *54*(3), 231-249.

Hinkel, K. M. (1997). Estimating seasonal values of thermal diffusivity in thawed and frozen soils using temperature time series. *Cold Regions Science and Technology, 26*(1), 1-15.

Hinkle D. E., W. Wiersma & S. G. Jurs (2003). Applied Statistics for the Behavioral Sciences (5th ed.). 756 p. Boston: Houghton Mifflin.

Hirota, T., Y. Iwata, M. Hayashi, S. Suzuki, T. Hamasaki, R. Sameshima & I. Takayabu (2006). Decreasing soil-frost depth and its relation to climate change in Tokachi, Hokkaido, Japan. *Journal of the Meteorological Society of Japan. Ser. II*, *84*(4), 821-833.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.

Hurrell, J. W., Kushnir, Y., Ottersen, G., & M. Visbeck (2003). An overview of the North Atlantic oscillation. The North Atlantic Oscillation: climatic significance and environmental impact, 134, 1-35.

Huttunen, L. & J. Soveri (1993) Luonnontilaisen roudan alueellinen ja ajallinen vaihtelu Suomessa. Vesi- ja ympäristöhallitus. 77 p.

Iijima, Y., A. N. Fedorov, H. Park, K. Suzuki, H. Yabuki, T. C. Maximov & T. Ohata (2010). Abrupt increases in soil temperatures following increased precipitation in a permafrost region, central Lena river basin, Russia. *Permafrost and Periglacial Processes, 21*(1), 30-41.

Iwata, Y., M. Hayashi & T. Hirota (2008). Comparison of snowmelt infiltration under different soil-freezing conditions influenced by snow cover. *Vadose Zone Journal, 7*(1), 79 p.

IPCC (2014): Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R. K. Pachauri and L. A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 p.

Jylhä, K., H. Tuomenvirta & K. Ruosteenoja (2004). Climate change projections for Finland during the 21st century. *Boreal Environment Research*, *9*(2), 127-152.

Jylhä, K., S. Fronzek, H. Tuomenvirta, T. Carter & K. Ruosteenoja (2008). Changes in frost, snow and Baltic Sea ice by the end of the twenty-first century based on climate model projections for Europe. *Climatic Change, 86*(3), 441-462.

Jylhä, K., K. Ruosteenoja, J. Räisänen, A. Venäläinen, H. Tuomenvirta, L. Ruokolainen, S. Saku & T. Seitola (2009). The changing climate in Finland: estimates for adaptation studies. ACCLIM project report 2009:4, 102 p.

Kellomäki, S., M. Maajärvi, H. Strandman, A. Kilpeläinen & H. Peltola (2010). Model computations on the climate change effects on snow cover, soil moisture and soil frost in the boreal conditions over Finland. *Silva Fennica, 44*(2).

Kersalo, J. & P. Pirinen (2009). *Suomen maakuntien ilmasto*. Raportteja 2009:8, Ilmatieteen laitos, Helsinki. 192 p.

Kim, Y., Kimball, J. S., McDonald, K. C., & J. Glassy (2011). Developing a global data record of daily landscape freeze/thaw status using satellite passive microwave remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(3), 949-960.

Kim, Y., J. S. Kimball, J. Glassy & J. Du (2017). An extended global earth system data record on daily landscape freeze–thaw status determined from satellite passive microwave remote sensing. *Earth System Science Data, 9*(1), 133-147.

Korhonen, K-H., Gardemeister, R. & M. Tammirinne (1974). *Geotekninen maaluokitus*. Valtion teknillinen tutkimuskeskus. Geotekniikan laboratorio, tiedonanto 14, 20 p.

Lean, J., & D. Rind (1996). The sun and climate. 7 p.

Lehtinen, M., P. Nurmi & T. Rämö (1998). *Suomen kallioperä: 3000 vuosimiljoonaa*. 375 p. Suomen geologinen seura.

Liu, J., S. Kang, T. Gong & A. Lu (2010). Growth of a high-elevation large inland lake, associated with climate change and permafrost degradation in Tibet. *Hydrology and Earth System Sciences, 14*(3), 481-489.

Longobardi, A. & P. Villani (2010). Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. *International journal of Climatology*, *30*(10), 1538-1546.

Luo, S., X. Fang, S. Lyu, Q. Jiang & J. Wang (2017). Interdecadal changes in the freeze depth and period of frozen soil on the three rivers source region in china from 1960 to 2014. *Advances in Meteorology, 2017*, 1-14.

Luoto, M. & J. Hjort (2005). Evaluation of current statistical approaches for predictive geomorphological mapping. *Geomorphology, 67*(3), 299-315.

Mitchell, J. F. (1989). The “greenhouse” effect and climate change. *Reviews of Geophysics*, *27*(1), 115-139.

Molenberghs, G. & G. Verbeke (2001). A review on linear mixed models for longitudinal data, possibly subject to dropout. *Statistical Modelling*, *1*(4), 235-269

Oechel, W. C., & G. L. Vourlitis (1994). The effects of climate charge on land—atmosphere feedbacks in arctic tundra regions. *Trends in Ecology & Evolution*, *9*(9), 324-329.

Orradottir, B., S. R. Archer, O. Arnalds, L. P. Wilding & T. L. Thurow (2008). Infiltration in Icelandic Andisols: the role of vegetation and soil frost. *Arctic, Antarctic, and Alpine Research*, *40*(2), 412-421.

Orvomaa M. (2015) Routa Suomessa 1981-2010: Suomen ympäristökeskus, <*http://harmosnow.eu/workshops/02.11.2015,%20Helsinki/Orvomaa_Pyry2015_Routa.pdf*> 1.4.2018

Osterkamp, T. E. (2007). Characteristics of the recent warming of permafrost in Alaska. *Journal of Geophysical Research - Earth Surface, 112*(F2), 1-10.

Peng, X., T. Zhang, B. Cao, Q. Wang, K. Wang, W. Shao & H. Guo (2016). Changes in freezing-thawing index and soil freeze depth over the Heihe River Basin, western China. *Arctic, Antarctic, and Alpine Research*, *48*(1), 161-176.

Pinheiro, J. C., Liu, C., & Y.N. Wu (2001). Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. *Journal of Computational and Graphical Statistics, 10*(2), 249-276.

Pirinen, P., H. Simola, J. Aalto, J. Kaukoranta, P. Karlsson & R. Ruuhela (2012). *Tilastoja suomen ilmastosta 1981 - 2010* Ilmatieteen laitos. 83 p.

Ronkainen, N. (2012). Suomen maalajien ominaisuuksia. Suomen ympäristökeskus. 57.

Ruosteenoja, K., K. Jylhä & H. Tuomenvirta (2005). *Climate scenarios for FINADAPT studies of climate change adaptation. FINADAPT Working Paper 15*. Finnish Environment Institute. 32 p.

Salmi, T., A. Määttä, P. Anttila, T. Ruoho-Airola & T. Amnell (2002). Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's slope estimates -the excel template application MAKESENS manual. Ilmatieteen laitos. 35.

Seppälä, M. (1999). Geomorphological aspects of road construction in a cold environment, Finland. *Geomorphology, 31*(1), 65-91.

Serreze, M. C., A. P. Barrett, J. C. Stroeve, D. N. Kindig & M. M Holland (2009). The emergence of surface-based Arctic amplification. *The Cryosphere*, *3*(1), 11-19.

Shiklomanov, N. I. (2012). Non-climatic factors and long-term, continental-scale changes in seasonally frozen ground. *Environmental Research Letters*, *7*(1).

Sinha T., & K. A. Cherkauer (2008). Time series analysis of soil freeze and thaw processes in Indiana. *Journal of Hydrometeorology, 9*(5), 936-950.

Solantie, R. (1998). Occurrence of unfrozen ground in Finland. *Geophysica*, *34*(3), 141-157.

Soveri, J. & M. Varjo (1977). *Roudan muodostumisesta ja esiintymisestä Suomessa vuosina 1955-1975.*Vesihallitus. 66 p.

Strahler, A. (2013). *Introducing physical geography.* (6th ed.). 641 p. New York, NY: Wiley.

Sutinen, R., P. Hänninen & A. Venäläinen (2008). Effect of mild winter events on soil water content beneath snowpack. *Cold Regions Science and Technology, 51*(1), 56-67.

Thomas, H. R., P. Cleall, Y. Li, C. Harris & M. Kern-Luetschg (2009). Modelling of cryogenic processes in permafrost and seasonally frozen soils. *Géotechnique, 59*(3), 173-184.

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research, 47*(1/2), 123-138.

Vatcheva, K. P., M. Lee, J. B. McCormick & M. H. Rahbar (2016). Multicollinearity in regression analyses conducted in epidemiologic studies. *Epidemiology (Sunnyvale, Calif.), 6*(2), 227 p.

Vaughan, D. G., J.C. Comiso, I. Allison, J. Carrasco, G. Kaser, R. Kwok, P. Mote, T. Murray, F. Paul, J. Ren, E. Rignot, O. Solomina, K. Steffen & T. Zhang (2013): Observations: cryosphere. *Climate change*, *2103*, 317-382.

Venäläinen, A., Tuomenvirta, H., Lahtinen, R., & Heikinheimo, M. (2001a). The influence of climate warming on soil frost on snow-free surfaces in Finland. *Climatic Change, 50*(1), 111-128.

Venäläinen, A. H. Tuomenvirta, M. Heikinheimo, S. Kellomäki, H. Peltola, H. Strandman & H. Väisänen (2001b). Impact of climate change on soil frost under snow cover in a forested landscape. *Climate Research, 17*(1), 63-72.

Wood, S. N. (2006). *Generalized additive models: an introduction with R*. 410 p. Chapman and Hall/CRC.

Yershov, E. D. (2004). *General geocryology*. 608 p. Cambridge university press.

Zhang, T., Barry, R. G., Knowles, K., Ling, F., & R.L Armstrong (2003). Distribution of seasonally and perennially frozen ground in the Northern Hemisphere. In *Proceedings of the 8th International Conference on Permafrost, 2*, 1289-1294.

Zhang, T. (2005). Influence of the seasonal snow cover on the ground thermal regime: An overview. *Reviews of Geophysics, 43*(4), 1-23.

Zhao, T., L. Zhang, L. Jiang, S. Zhao, L. Chai & R. Jin (2011). A new soil freeze/thaw discriminant algorithm using AMSR-E passive microwave imagery. *Hydrological Processes, 25*(11), 1704-1716.

Zuur, A. F., E. N Ieno, N. J Walker, A. A Saveliev & G. M Smith (2009). GLM and GAM for count data. In: *Mixed effects models and extensions in ecology with R*. Springer, New York, NY. 209-243.

Öquist, M. G., & Laudon (2008). Winter soil frost conditions in boreal forests control growing season soil CO2 concentration and its atmospheric exchange. *Global Change Biology, 14*(12), 2839-2847.

Digital references

"All About Frozen Ground." National Snow and Ice Data Center. Accessed 1 January 2018. <*/cryosphere/frozenground/index.html*>

Hydrologisen seurannan kenttätöiden toimintakäsikirja v1.2 Suomen ympäristökeskus. 97 p. 26.10.2017

Ilmasto-opas. Suomen muuttuva ilmasto. Ilmatieteen laitos <*https://ilmasto-opas.fi/fi/ilmastonmuutos/suomen-muuttuva-ilmasto/-/artikkeli/1c8d317b-5e65-4146-acda-f7171a0304e1/nykyinen-ilmasto-30-vuoden-keskiarvot.html*> 1.4.2018

Kaivannaistietoa kaikille. Geologian tutkimuskeskus 1.4.2018 <*www.kaiva.fi/geologia*>

North Atlantic Oscillation (NAO). National Oceanic and Atmospheric Administriation. 10.1.2012 <*http://www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml*>

National Land Survey of Finland. 2010. <*https://epsg.io/2393*> 1.4.2018

Solar radiation basics. U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy. 1.4.2018 <*https://www.energy.gov/eere/solar/articles/solar-radiation-basics*>

Data authors

Finnish Meteorological Insitute, Daily Mean Temperature, 10km, 1961-2014, GeoTIFF, CSC - IT Center for Science Ltd, 2016-06-29T12:52:51+00:00, urn:nbn:fi:csc-kata00001000000000000661

Finnish Meteorological Insitute, Monthly Mean Temperature, 10km, 1961-2014, GeoTIFF, CSC - IT Center for Science Ltd, 2016-06-29T12:52:51+00:00, urn:nbn:fi:csc-kata00001000000000000679

Finnish Meteorological Insitute, Monthly Precipitation Sum, 10km, 1961-2014, GeoTIFF, CSC - IT Center for Science Ltd, 2016-06-29T12:52:51+00:00, urn:nbn:fi:csc-kata00001000000000000685

Finnish Meteorological Insitute, Daily global radiation, 10km, 1961-2014, GeoTIFF, CSC – IT Center for Science Ltd, 2016-06-29T12:52:51+00:00, urn:nbn:fi:csc-kata00001000000000000748

National Land Survey of Finland, Elevation model 2000, 25 m x 25 m, TIFF, CSC - IT Center for Science Ltd, 2015-10-08T00:00:00-00:00, urn:nbn:fi:csc-kata00001000000000000319

Mann-Kendall Test, & Sen's Slope Estimates for the Trend of Annual Data. *Makesens 1.0*: Salmi, T., A. Määttä, P. Anttila, T. Ruoho-Airola & T. Amnell (2002). Detecting trends of annual values of atmospheric pollutants by the Mann-Mendall test and Sen's slope estimates -the excel template application MAKESENS. <*https://en.ilmatieteenlaitos.fi/makesens*> 1.4.2018

11. Appendix

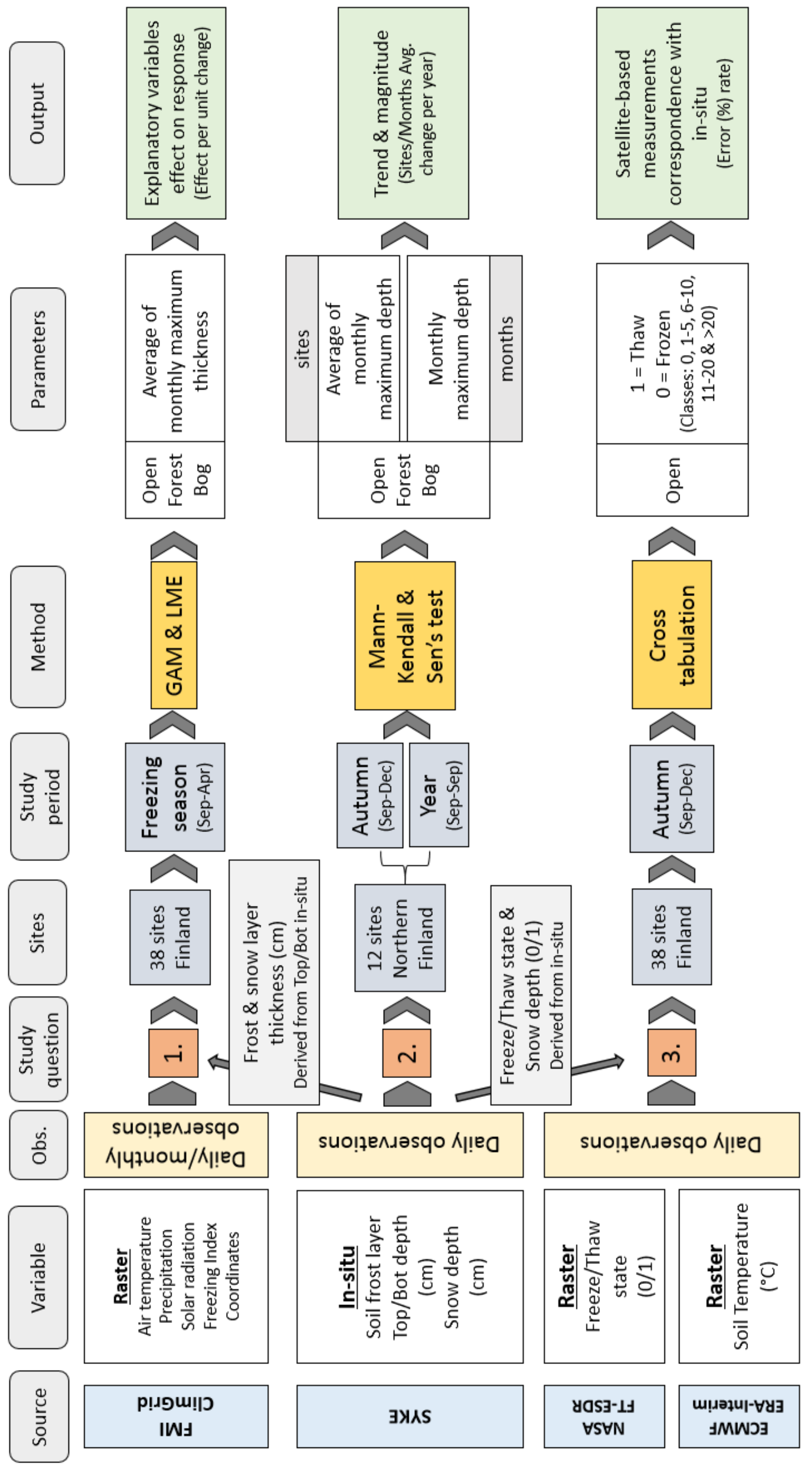
Appendix 1. Soil type classification based on grain diameter (Aaltonen et al. 1949; Korhonen et al. 1975) Modified by author.

Geotechnical soil type classification	Grain diameter (mm)	Engrineering soil type classification
Boulders	>600	Boulders
	600-200	Rocks
Rocks	200-60	
	60-20	Gravel
Gravel	20-6.	
	6-2.	
Sand	2-0,6	Sand
	0,6-0,2	
Finesand	0,2-0,06	
	0,06-0,02	Silt
Silt	0,02-0,006	
	0,006-0,002	
Clay	<0,002	Clay

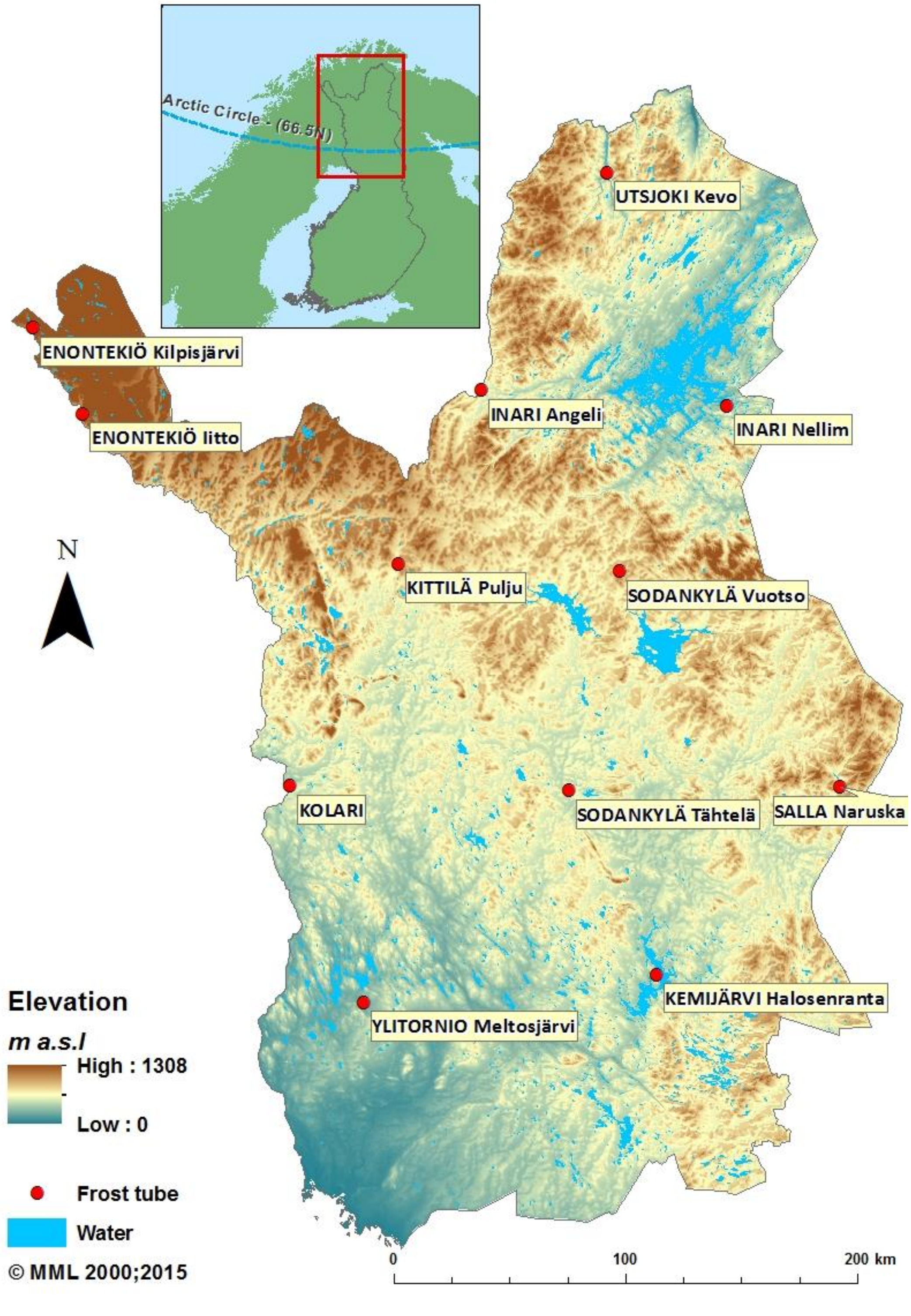
Appendix 2. Summary of all soil types and site specific descriptions and soil types of northern Finland frost tubes sites (Aaltonen, et al. 1949; Huttunen & Soveri 1993). Modified by author.

Soil type		
Finnish code	**English name**	**Finnish name**
Sr	Gravel	Sora
Mr	Moraine	Moreeni
SrMr	Gravel moraine	Soramoreeni
HkMr	Sand moraine	Hiekkamoreeni
HtMr	Finesand moraine	Hietamoreeni
HsMr	Silt moraine	Hiesumoreeni
SMr	Clay moraine	Savimoreeni
Hk	Sand soil	Hiekkamaa
KHk	Coarse sand	Karkea hiekka
HHk	Sand	Hieno hiekka
Ht	Finesand soil	Hietamaa
KHt	Finesand	Karkea hieta
HHt	Finer finesand	Hieno hieta
Hs	Silt	Hiesu
Sa	Clay soil	Savimaa
Tv	Peat	Turve
Hu	Humus	Humus

Site	**Open area**	**Forest**
Enontekiö litto	Tv	HuHsHt
Enontekiö Kilpisjärvi	-	Sa
Inari Angeli	HHk	HtHk
Inari Nellim	HHk	Hk
Kemijärvi	HkMr	HkMr
Kittilä	Sa	HkSr
Kolari	KHt	KHt
Salla	HkHt	HkHt
Sodankylä Tähtelä	SrHk	SrHk
Sodankylä Vuotso	HkSr	Hk
Utsjoki	-	Hk
Ylitornio	Ht	Ht



Appendix 3. Study settings of study questions: 1) Which variables affect the overall thickness of soil frost layer and its changes most? 2) Are there trends or major changes in the past 30 years on soil frost time series in northern Finland? 3) Correspondence of Freeze/Thaw-Earth System Data Record and ERA Interim soil temperature dataset with in situ soil frost observations in Finland.



Appendix 4. Used frost tube sites in Northern Finland trend analysis. Coordinate system: KKJ/Finland Uniform Coordinate System.

Appendix 5. Spearman's correlation matrixes for response-explanatory variables for three different type of frost tube site. Over 0.4 absolute correlations highlighted.

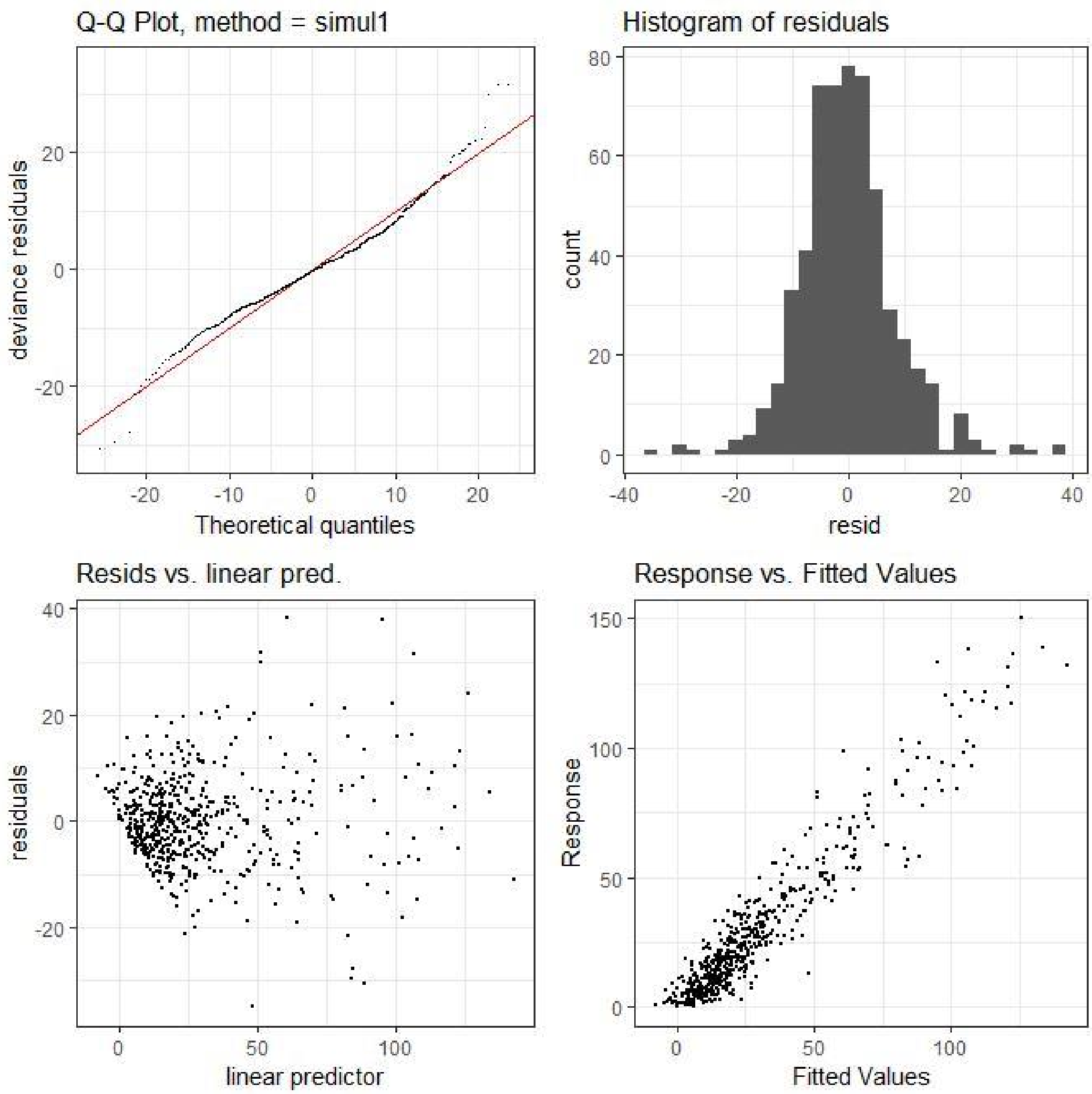
OPEN AREA	AVE.MAX.Frostlaye	AVE.MAX.Snow	AVE.NAO	AVE.Prec	AVE.Rad	AVE.Temp	Freezing.Index	coord.east	coord.north
AVE.MAX.Frostlayer	1								
AVE.MAX.Snow	0.163	1							
AVE.NAO	-0.030	-0.101	1						
AVE.Prec	**-0.575**	-0.218	0.042	1					
AVE.Rad	**-0.448**	**-0.639**	-0.141	0.410	1				
AVE.Temp	**-0.639**	**-0.691**	0.193	**0.594**	**0.699**	1			
Freezing.Index	0.041	0.071	-0.321	-0.096	-0.014	-0.228	1		
coord.east	-0.017	**0.588**	-0.031	-0.025	-0.374	**-0.407**	0.168	1	
coord.north	**0.610**	**0.652**	-0.007	**-0.581**	**-0.878**	**-0.840**	0.086	0.341	1

FOREST	AVE.MAX.Frostlayer	AVE.MAX.Snow	AVE.NAO	AVE.Prec	AVE.Rad	AVE.Temp	Freezing.Index	coord.east	coord.north
AVE.MAX.Frostlayer	1								
AVE.MAX.Snow	-0.135	1							
AVE.NAO	-0.169	-0.083	1						
AVE.Prec	**-0.584**	-0.027	0.199	1					
AVE.Rad	**-0.508**	0.125	-0.145	0.283	1				
AVE.Temp	**-0.761**	0.117	0.204	**0.591**	**0.686**	1			
Freezing.Index	0.036	0.164	-0.307	-0.062	0.070	-0.098	1		
coord.east	0.218	-0.313	-0.000	-0.035	-0.378	**-0.441**	0.135	1	
coord.north	**0.658**	-0.142	0.001	**-0.510**	**-0.791**	**-0.855**	-0.061	0.362	1

BOG	AVE.MAX.Frostlaye	AVE.MAX.Snow	AVE.NAO	AVE.Prec	AVE.Rad	AVE.Temp	Freezing.Index	coord.east	coord.north
AVE.MAX.Frostlayer	1								
AVE.MAX.Snow	0.227	1							
AVE.NAO	-0.013	-0.051	1						
AVE.Prec	-0.177	-0.113	0.161	1					
AVE.Rad	-0.076	-0.274	-0.133	0.342	1				
AVE.Temp	-0.192	-0.248	0.183	**0.600**	**0.704**	1			
Freezing.Index	-0.032	0.012	-0.261	-0.094	0.005	-0.108	1		
coord.east	-0.001	0.105	-0.004	-0.033	-0.193	-0.251	0.182	1	
coord.north	0.112	0.227	0.008	**-0.562**	**-0.789**	**-0.865**	-0.032	0.195	1

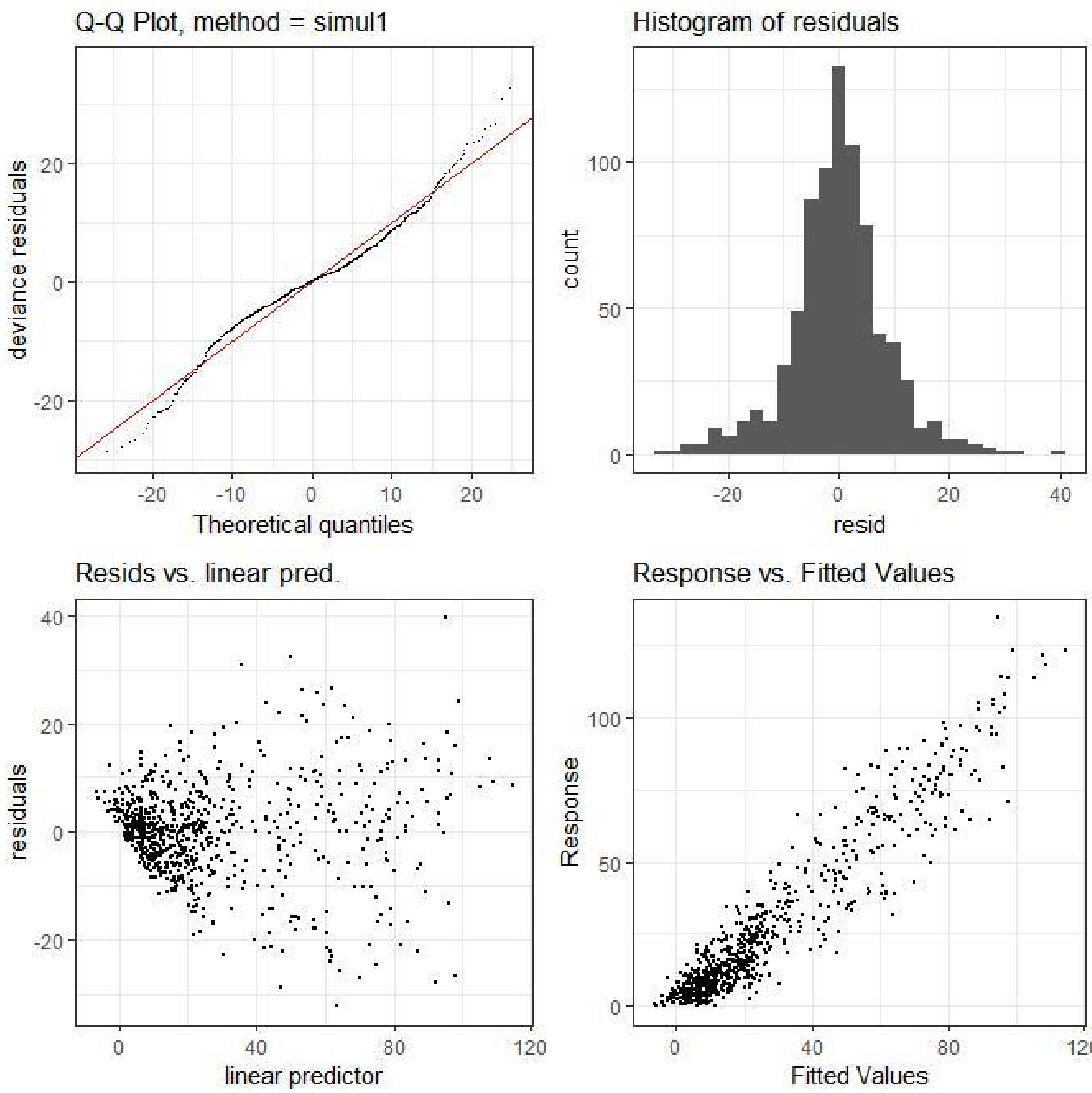
Appendix 6. Residual plots of *gam.check*. (a) open area, (b) forest, (c) bog.

(a)



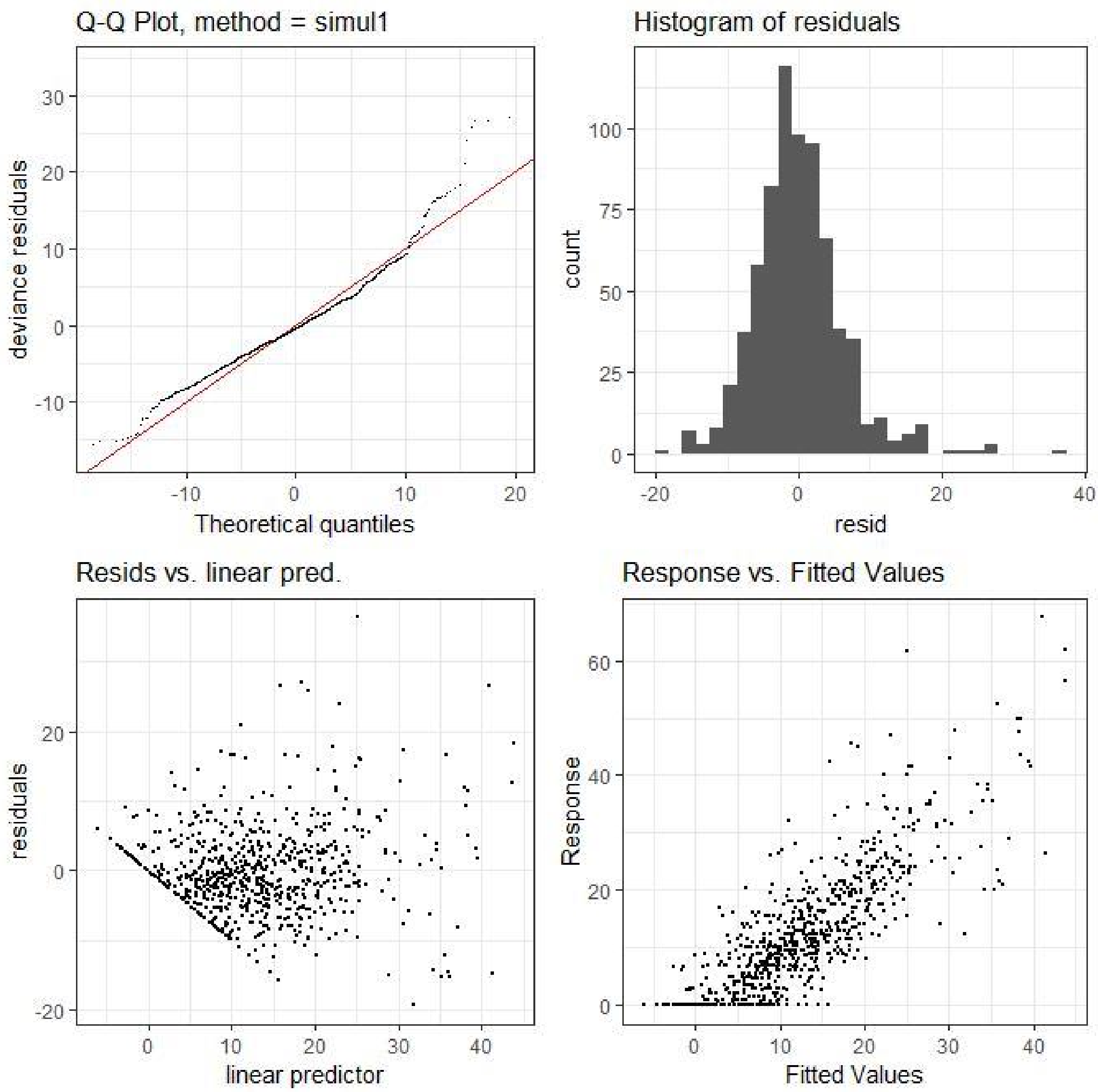
Appendix 6 continues.

(b)



Appendix 6 continues.

(c)



Appendix 7. Contingency table based on soil frost in-situ observations and NASA's ESDR landscape freeze/thaw prediction. Data is divided into four class (0-5, 6-10, 11-20 and >20 cm) based on soil frost and snow depth observations.

				FT-ESDR														
				SNOW (cm)														
				0			1-5			6-10			11-20			>20		
SYKE in-situ	SOIL FROST (cm)			Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall
		0	Thaw	104	507	611	19	40	59	21	16	37	25	29	54	52	11	63
			Tab %	17	83		32	68		57	43		46	54		83	17	
			Totall	104	507	**611**	19	40	**59**	21	16	**37**	25	29	**54**	52	11	**63**
		1-5	Frozen	36	111	147	85	88	173	65	41	106	111	40	151	182	22	204
			Tab %	24	76		49	51		61	39		74	26		89	11	
			Totall	36	111	**147**	85	88	**173**	65	41	**106**	111	40	**151**	182	22	**204**
		6-10	Frozen	22	40	62	63	41	104	52	27	79	87	37	124	159	32	191
			Tab %	35	65		61	39		66	34		70	30		83	17	
			Totall	22	40	**62**	63	41	**104**	52	27	**79**	87	37	**124**	159	32	**191**
		11-20	Frozen	15	36	51	41	32	73	42	26	68	121	19	140	174	26	200
			Tab %	29	71		56	44		62	38		86	14		87	13	
			Totall	15	36	**51**	41	32	**73**	42	26	**68**	121	19	**140**	174	26	**200**
		>20	Frozen	4	6	10	46	19	65	47	16	63	164	31	195	435	21	456
			Tab %	40	60		71	29		75	25		84	16		95	5	
			Totall	4	6	**10**	46	19	**65**	47	16	**63**	164	31	**195**	435	21	**456**

Appendix 8. Contingency table based on soil frost in-situ observations and ECMWF's soil temperature predictions on 7 cm below the ground surface. Data is divided into four class (0-5, 6-10, 11-20 and >20 cm) based on soil frost and snow depth observations.

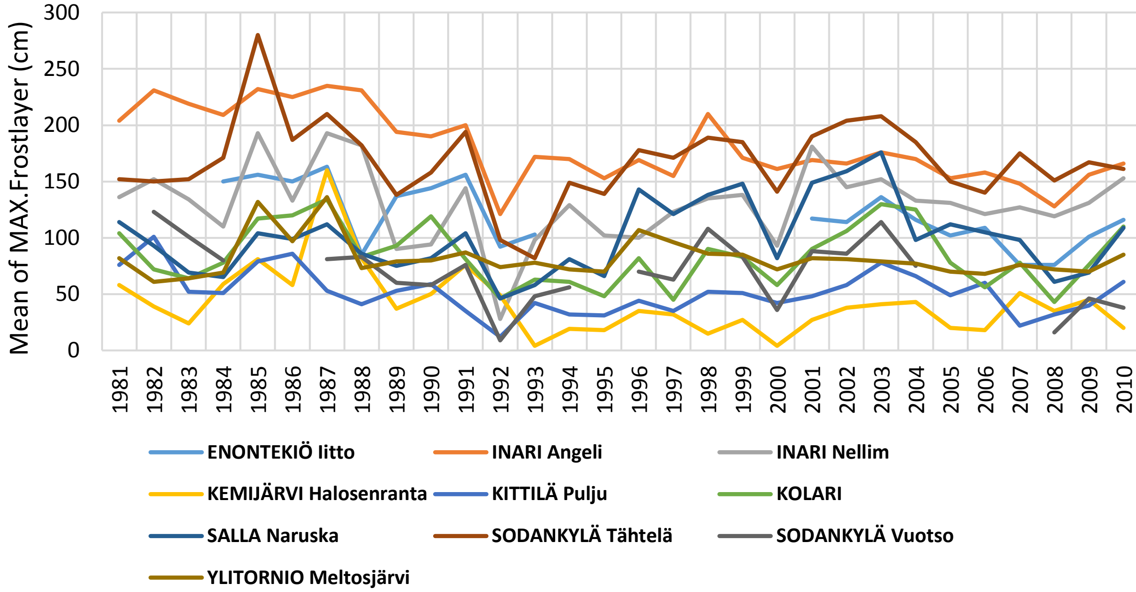
				ERA Interim														
				SNOW (cm)														
				0			1-5			6-10			11-20			>20		
SYKE in-situ	SOIL FROST (cm)			Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall
		0	Thaw	92	519	611	13	46	59	8	29	37	22	32	54	47	16	63
			Tab %	15	85		22	78		22	78		41	59		75	25	
			Totall	92	519	**611**	13	46	**59**	8	29	**37**	22	32	**54**	47	16	**63**
		1-5	Frozen	47	100	147	92	81	173	67	39	106	112	39	151	179	25	204
			Tab %	32	68		53	47		63	37		74	26		88	12	
			Totall	47	100	**147**	92	81	**173**	67	39	**106**	112	39	**151**	179	25	**204**
		6-10	Frozen	23	39	62	75	29	104	59	20	79	105	19	124	176	15	191
			Tab %	37	63		72	28		75	25		85	15		92	8	
			Totall	23	39	**62**	75	29	**104**	59	20	**79**	105	19	**124**	176	15	**191**
		11-20	Frozen	29	22	51	56	17	73	55	13	68	123	17	140	176	15	191
			Tab %	57	43		77	23		81	19		88	12		92	8	
			Totall	29	22	**51**	56	17	**73**	55	13	**68**	123	17	**140**	176	15	**191**
		>20	Frozen	5	5	10	60	5	65	53	10	63	180	15	195	439	17	456
			Tab %	50	50		92	8		84	16		92	8		96	4	
			Totall	5	5	**10**	60	5	**65**	53	10	**63**	180	15	**195**	439	17	**456**

Appendix 9. Contingency table based on NASA's ESDR landscape freeze/thaw and ECMWF's soil temperature predictions. Data is divided into four class (0-5, 6-10, 11-20 and >20 cm) based on soil frost and snow depth observations.

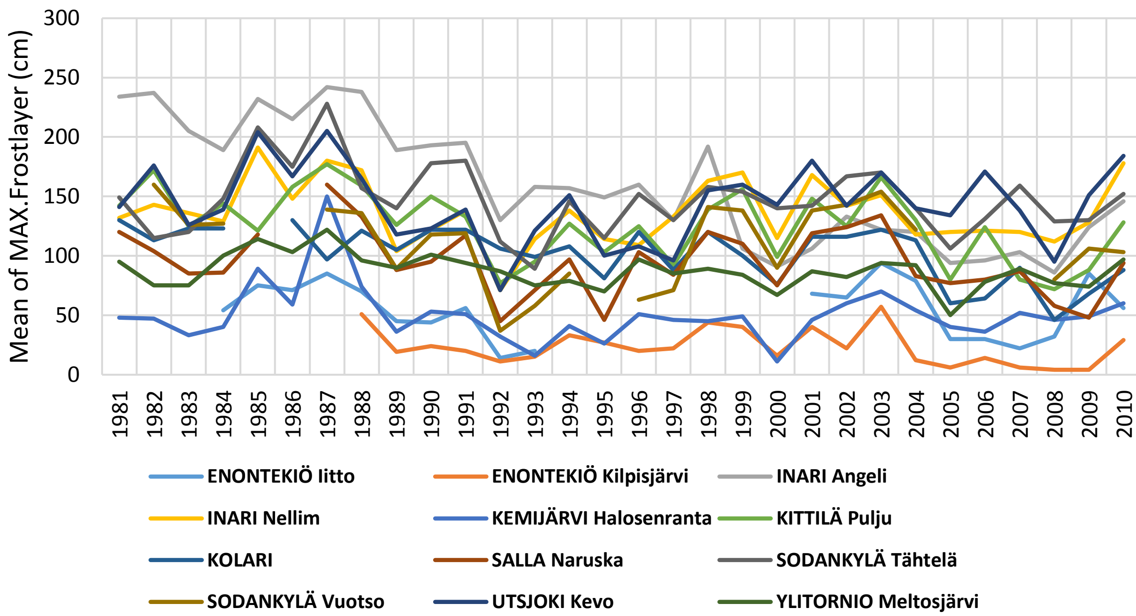
			FT-ESDR														
			SNOW (cm)														
			0			1-5			6-10			11-20			>20		
ERA Interim	SOIL FROST (cm)		Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall	Frozen	Thaw	Totall
	0-5	**Frozen**	38	66	104	5	14	19	47	28	75	94	40	134	202	24	226
		Tab %	6	11		8	24		33	20		46	20		76	9	
		Thaw	54	453	507	8	32	40	39	29	68	42	29	71	32	9	41
		Tab %	9	74		14	54		27	20		20	14		12	3	
		Totall	92	519	**611**	13	46	**59**	86	57	**143**	136	69	**205**	234	33	**267**
	6-10	**Frozen**	8	28	36	49	36	85	41	18	59	77	28	105	151	25	176
		Tab %	5	19		28	21		52	23		62	23		79	13	
		Thaw	39	72	111	43	45	88	11	9	20	10	9	19	8	7	15
		Tab %	27	49		25	26		14	11		8	7		4	4	
		Totall	47	100	**147**	92	81	**173**	52	27	**79**	87	37	**124**	159	32	**191**
	11-20	**Frozen**	6	16	22	47	16	63	35	20	55	108	15	123	161	24	185
		Tab %	10	26		45	15		51	29		77	11		80	12	
		Thaw	17	23	40	28	13	41	7	6	13	13	4	17	13	2	15
		Tab %	27	37		27	12		10	9		9	3		6	1	
		Totall	23	39	**62**	75	29	**104**	42	26	**68**	121	19	**140**	174	26	**200**
	>20	**Frozen**	8	7	15	34	7	41	41	12	53	152	28	180	421	18	439
		Tab %	16	14		47	10		65	19		78	14		92	4	
		Thaw	21	15	36	22	10	32	6	4	10	12	3	15	14	3	17
		Tab %	41	29		30	14		10	6		6	2		3	1	
		Totall	29	22	**51**	56	17	**73**	47	16	**63**	164	31	**195**	435	21	**456**

Appendix 10. Soil frost maximum depth in northern Finland 1981-2010.

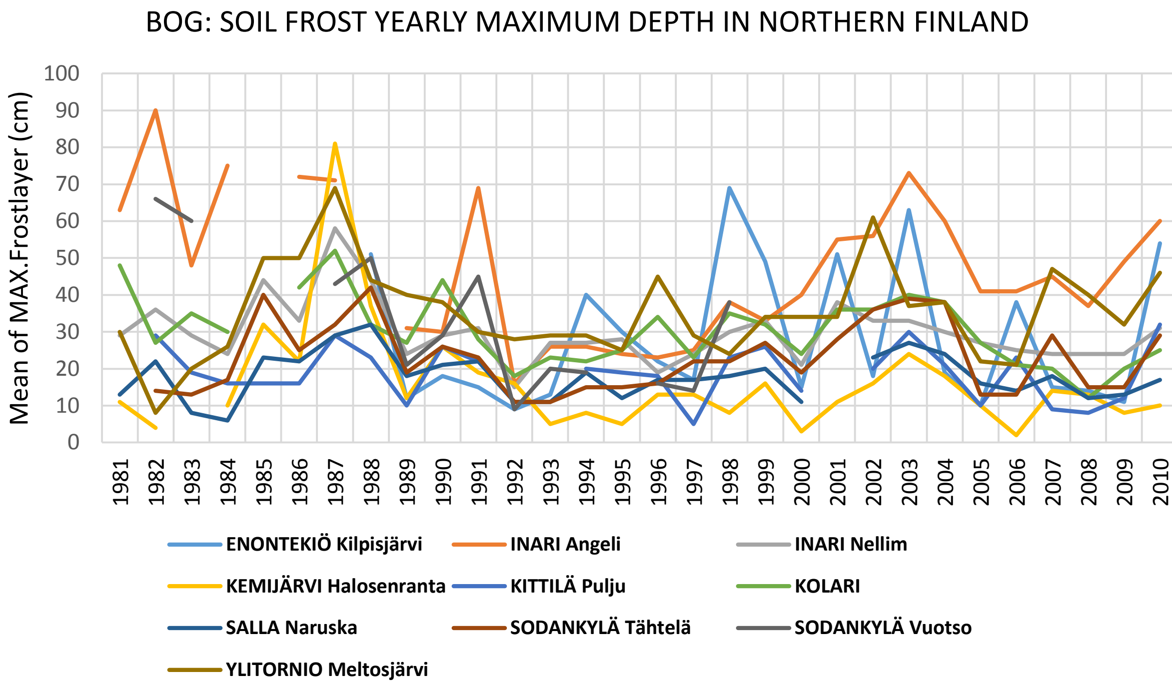
OPEN AREA: SOIL FROST YEARLY MAXIMUM DEPTH IN NORTHERN FINLAND

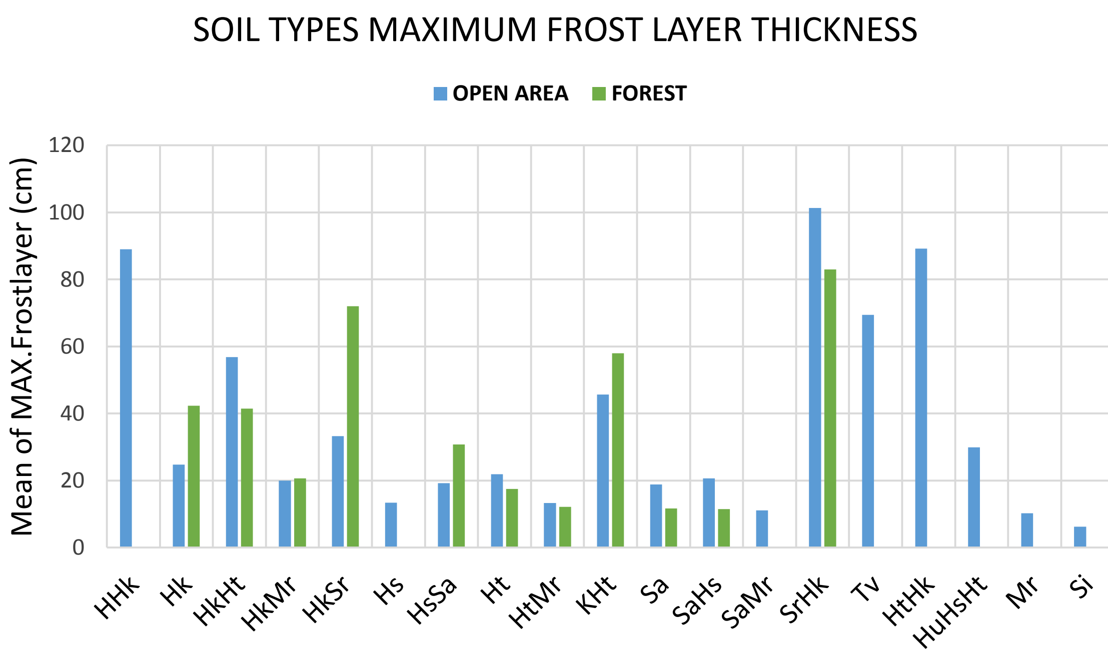


FOREST: SOIL FROST YEARLY MAXIMUM DEPTH IN NORTHERN FINLAND



Appendix 10 continues.





Appendix 11. Soil types mean maximum of soil frost layer thickness in open and forest areas (Ronkainen 2012). Description in Appendix 2.